ACTA SCIENTIARUM MATHEMATICARUM

TOMUS XII.

LEOPOLDO FEJÉR ET FREDERICO RIESZ
LXX ANNOS NATIS DEDICATUS.

PARS B.

SZEGED, 1950.

TOMUM IUBILAREM
ADIUVANTE ACADEMIA SCIENTIARUM HUNGARICA
EDIDERUNT
INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS
ET SOCIETAS MATHEMATICA DE IOHANNE BOLYAI NOMINATA

ACTA SCIENTIARUM MATHEMATICARUM

12. KÖTET

FEJÉR LIPÓT ÉS RIESZ FRIGYES

70. SZÜLETÉSNAPJÁRA.

B. RÉSZ

S Z E G E D, 1950.

AZ ÜNNEPI KÖTETET
A MAGYAR TUDOMÁNYOS AKADÉMIA TÁMOGATÁSÁVAL
KIADTA
A SZEGEDI TUDOMÁNYEGYETEM BOLYAI-INTÉZETE
ÉS A BOLYAI JÁNOS MATEMATIKAI TÁRSULAT

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS XII.

LEOPOLDO FEJÉR ET FREDERICO RIESZ

LXX ANNOS NATIS DEDICATUS.

PARS B.

SZEGED, 1950.

TOMUM IUBILAREM
ADIUVANTE ACADEMIA SCIENTIARUM HUNGARICA
EDIDERUNT
INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS
ET SOCIETAS MATHEMATICA DE IOHANNE BOLYAI NOMINATA

ACTA SCIENTIARUM MATHEMATICARUM

12. KÖTET

FEJÉR LIPÓT ÉS RIESZ FRIGYES

70. SZÜLETÉSNAPJÁRA.

B. RÉSZ

SZEGED, 1950.

AZ ÜNNEPI KÖTETET
A MAGYAR TUDOMÁNYOS AKADÉMIA TÁMOGATÁSÁVAL
KIADTA
A SZEGEDI TUDOMÁNYEGYETEM BOLYAI-INTÉZETE
ÉS A BOLYAI JÁNOS MATEMATIKAI TÁRSULAT

Lattice points and Fourier expansions.

By S. BOCHNER in Princeton, N. J., and K. CHANDRASEKHARAN in Bombay.

1. Introduction.

We have recently given [4][1]) a new line of reasoning for proving HARDY's identity [8] in the theory of lattice points in a circle, and for the related convergence theorems of HARDY, LANDAU [8, 9], WALFISZ [12, 13], OPPENHEIM [11], WILTON [15, 16], DIXON and FERRAR [7]. We employed a general summability-theorem, due to BOCHNER [3, Th. 1], for partial derivatives of multiple Fourier series, and we combined it with a theorem of ANANDA-RAU [1] on scales of Riesz summability for general Dirichlet series in which assumptions on the magnitude of the coefficients are made explicitly.

In the present paper we will throw the part due to ANANDA-RAU into the differentiability-theorem itself, thus obtaining a much broader theorem on multiple Fourier series in general, from which to deduce the particular lattice-point conclusions by much shorter steps. Actually in § 3 we will first have a relatively simple version of the general differentiability theorem sufficient for the lattice-point conclusions envisaged, and afterwards, in § 5 and § 6, we will enlarge on the differentiability theorem for its own sake. This will bring out its similarity with a criterion of CHANDRASEKHARAN and MINAKSHISUNDARAM [6, Th. 4.1] which was the first attempt towards extending, from one to several variables, a convergence-test for Fourier series due to HARDY and LITTLEWOOD [10] in which the order of magnitude of the Fourier coefficients is prescribed; and it will also throw further light on the entire problem of localization of convergence and summability for Fourier series in general [2]; the latter problem is more delicate for multiple series than for simple series, and rather more delicate for formal (partial) derivatives of a series than for the original series proper, and the present paper may also be viewed as a further contribution towards managing this problem in some of its aspects.

---

[1]) Numbers in brackets [ ] refer to the bibliography placed at the end of the paper.

2. Notations and Definitions.

Let $f(x)=f(x_1,\ldots,x_k)$ be periodic in each variable with period 2π, and Lebesgue integrable in (x). It has then a Fourier expansion which we indicate by writing

$$f(x_1,\ldots,x_k)\sim\sum\ldots\sum a_{n_1\ldots n_k}e^{i(n_1x_1+\ldots+n_kx_k)}.$$

Let

$$A_n(x)=\sum_{n_1^2+\ldots+n_k^2=n}a_{n_1\ldots n_k}e^{i(n_1x_1+\ldots+n_kx_k)}$$

and

$$S_n(x)=\sum_0^n A_r(x).$$

Define for $l>0$

$$S^l(x_1,\ldots,x_k;R)\equiv S^l(x;R)\equiv S^l(R)=$$

$$=\sum_{n_1^2+\ldots+n_k^2\leqq R^2}\{R^2-(n_1^2+\ldots+n_k^2)\}^l a_{n_1\ldots n_k}e^{i(n_1x_1+\ldots+n_kx_k)}=$$

$$=\sum_{r=0}^n(R^2-r)^l A_r(x)=2l\int_0^R(R^2-u^2)^{l-1}S(u)u\,du,$$

where

$$n=[R]\quad\text{and}\quad S(R)=S^0(R)=S_n(x).$$

Let

$$T^l(R)=S^l(R)R^{-2l}.$$

Define

$$f(x,t)=\frac{\Gamma(k/2)}{2(\pi)^{k/2}}\int_\sigma f(x_1+t\xi_1,\ldots,x_k+t\xi_k)\,d\sigma_\xi$$

where σ denotes the unit-sphere $\xi_1^2+\ldots+\xi_k^2=1$ and $d\sigma_\xi$ its $(k-1)$-dimensional volume-element.

3. A Convergence Theorem.

We shall first state a few lemmas, which are needed for the proof of our theorem.

Lemma 1. *Suppose that*

$$\frac{a_n}{l_n-l_{n-1}}=O(l_n^\alpha),$$

where $\{l_n\}$ is a strictly increasing sequence of positive numbers diverging to ∞, and suppose that $\sum a_n l_n^{-\gamma}$ is summable by Riesz's means of type l_n and of order r, briefly: summable (l,r), γ being real. Let $0\leqq s<r$. Then $\sum a_n l_n^{-\sigma}$ is summable (l,s) for

$$\sigma>\frac{(\alpha+1)(r-s)+\gamma(s+1)}{r+1}.$$

This has been proved by ANANDA-RAU [1, Th. 7], and if we choose $l_n = n$, $\gamma = 0$, it reduces to the following

Lemma 1 A. *If* $\sum a_n$ *is summable* (n, r) *and* $a_n = O(n^\alpha)$ *and* $0 \leqq s < r$, *then* $\sum a_n n^{-\sigma}$ *is summable* (n, s) *for*

$$\sigma > \frac{(\alpha+1)(r-s)}{r+1}.$$

Lemma 2. *If* $f(x) = f(x_1, \ldots, x_k)$ *is a periodic function of class L, or an almost periodic function of Stepanoff class, and*

$$f(x) \sim \sum_n a(n)\, e^{i\Lambda(n,x)} \tag{3.1}$$

where $\Lambda(n, x)$ *denotes* $n_1x_1 + \ldots + n_kx_k$, *and* $a(n) = a(n_1, \ldots, n_k)$ *is the Fourier coefficient, and* $D^q(n_1, \ldots, n_k)$ *is, for any non-negative integer q, a homogeneous polynomial of total degree q in* $n_1, \ldots, n_k$, *then*

(i) *the operator*

$$D_x^q = D^q\left(\frac{\partial}{i\partial x_1}, \ldots, \frac{\partial}{i\partial x_k}\right)$$

applies to the almost periodic function

$$T_R^\delta(x) = \sum_{|n| \leqq R}\left(1 - \frac{|n|^2}{R^2}\right)^\delta a(n)\, e^{i\Lambda(n,x)}$$

and the resulting function is almost periodic;

(ii) $$D_x^q T_R^\delta(x) = \sum_{|n| \leqq R}\left(1 - \frac{|n|^2}{R^2}\right)^\delta a(n)\, D^q(n) e^{i\Lambda(n,x)};$$

(iii) *for every x at which the condition*

$$\int_0^t |f_x(t)|\, t^{k-1-q}\, dt = o(t^k)$$

is satisfied, we have

$$\lim_{R \to \infty} D_x^q T_R^\delta(x) = 0 \tag{3.2}$$

for $\delta > \dfrac{k-1}{2} + q$.

This has been proved by BOCHNER [2, Th. I].

Lemma 3. *If* $k \geqq 1$, $0 \leqq n < \infty$, *and if the numbers* $a_{n_1 \ldots n_k}$ *are arbitrarily given for* $0 \leqq n_1 \leqq n, \ldots, 0 \leqq n_k \leqq n$, *then there exists an exponential polynomial*

$$P(x_1, \ldots, x_k) = \sum_{\lambda_1 = 0}^{n} \cdots \sum_{\lambda_k = 0}^{n} \gamma_{\lambda_1 \ldots \lambda_k} e^{i(\lambda_1 x_1 + \ldots + \lambda_k x_k)}$$

such that at the origin

$$\left(\frac{\partial^{n_1+\dots+n_k} P}{\partial x_1^{n_1}\dots\partial x_k^{n_k}}\right)_{x=(0)} = a_{n_1\dots n_k}.$$

Proof. Obviously, if arbitrary numbers $b_{n_1\dots n_k}$, $0 \leqq n_j \leqq n$, $(j=1,\dots,k)$ are prescribed, then there exists an (ordinary) polynomial

$$Q(z_1,\dots,z_k) = \sum_{\mu_1=0}^{n} \dots \sum_{\mu_k=0}^{n} \delta_{\mu_1\dots\mu_k} z_1^{\mu_1}\dots z_k^{\mu_k}$$

such that

$$\left(\frac{\partial^{n_1+\dots+n_k} Q}{\partial z_1^{n_1}\dots\partial z_k^{n_k}}\right)_{z=0} = b_{n_1\dots n_k},$$

namely, $b_{n_1\dots n_k} = n_1!\dots n_k!\,\delta_{n_1\dots n_k}$. Now consider the transformation of variables

$$z_1 = e^{ix_1} - 1, \dots, z_k = e^{ix_k} - 1.$$

Obviously it transforms a $P(x)$ into a $Q(z)$ and conversely, under preservation of n, and for prescribed values $a_{n_1\dots n_k}$ this leads to values $b_{n_1\dots n_k}$ by ordinary rules of differentiation of a function of functions, and inversely from the b's to the a's, and hence the lemma.

Lemma 4. *If $f(x)$ is a periodic or almost periodic function* (3. 1), *and if in a neighborhood of the point $x=x_0$ the function has continuous derivatives of total order $\leqq q$, then at $x=x_0$ we have for $\delta > \frac{k-1}{2}+q$:*

$$\lim_{R\to\infty} [D_x^q T_R^\delta(x) - D_x^q f(x)] = 0.$$

Proof. The conclusion is obviously trivial for an *exponential* polynomial $P(x)$. In general we put, by lemma 3, $f(x) = P(x) + f^1(x)$ where for $f^1(x)$ all partial derivatives of total order $\leqq q$ are zero at the point $x=x_0$. But $f^1(x)$ has also continuous derivatives of order q in the neighborhood of x_0. From this it follows easily that $f^1(x)$ satisfies assumption (iii) of lemma 2 and hence (3. 2) follows.

Remark. The "modification" referred to in lemma 6 of our previous paper [4, p. 241] is made explicit now; even there, differentiability has to be assumed in a neighborhood of the point in question.

Theorem 1. *If $f(x)$ is defined as in § 2, and if*

(3. 3) $$A_n = O(n^\alpha)$$

then at every point x in a neighborhood of which $f(x)$ possesses partial derivatives of all orders, the series $\sum A_n n^h$ is summable (n, δ) for $\delta \geqq 0$, and $\delta > 2\alpha + 1 + 2h$.

Proof. By lemma 4, if we choose D as the Laplace operator

$$\left(\frac{\partial^2 f}{\partial x_1^2}+\dots+\frac{\partial^2 f}{\partial x_k^2}\right)$$

and apply it q times to the function f, where q is a non-negative integer, we obtain that $\sum A_n n^q$ is summable (n, δ) for $\delta > \frac{k-1}{2}+2q$. Since $A_n n^q = O(n^{\alpha+q})$, it follows by lemma 1 A, that $\sum A_n n^h$ is summable (n, η) for $\eta \geqq 0$, and

$$q-h > \frac{(\alpha+q+1)(\delta-\eta)}{\delta+1}$$

or

$$\eta > \delta - \frac{(q-h)(\delta+1)}{\alpha+q+1}.$$

Since δ may be any number greater than $\frac{k-1}{2}+2q$, this implies that any

$$\eta > \frac{\left(\frac{k-1}{2}\right)(\alpha+1+h)+h+2q\left(\alpha+\frac{1}{2}+h\right)}{\alpha+q+1}$$

is admissible. Given k, α, h since q may be chosen as large as we please, the theorem will be true for $\eta > 2\alpha+1+2h$.

Remarks. It should be noticed that there is no restriction on α. However, if

$$a_{n_1 \dots n_k} = O\left(\frac{1}{(n_1^2+\dots+n_k^2)^{k/2}}\right) \tag{3.4}$$

then at every point of mean-continuity (etc.) we have convergence of $\sum A_n$. See [6, p. 741]. The significance of the theorem is that even though only something less than (3.4) is satisfied, a stronger hypothesis on the function than continuity will still lead to summability, and, in special cases, to convergence. We will show in the next section how the above theorem is entirely adequate to obtain the most complete results on the summation of certain series of Bessel functions occurring in the theory of lattice points.

4. Application to summations over lattice points.

Let $r_k(n) = \sum_{n_1^2+\dots+n_k^2=n} 1$ for integral values of n_k, representation of n which differ only in sign or order being counted as distinct. Let

$$R_k(x) = \sum_{n \leqq x}{}' r_k(n)$$

the last term $r_k(x)$ in the sum being replaced by $\frac{1}{2} r_k(x)$ if x is an integer.

Then it is known that $R_k(x)$ can be "represented" as a series of Bessel functions; in particular, if $k=2$, we have HARDY's identity [8]: if x is non-integral,

$$R_2(x)=\pi x-x^{1/2}\sum\frac{r_2(n)J_1(2\pi\sqrt{nx})}{n^{1/2}}. \tag{4.1}$$

Here J_1 stands for the Bessel function of order 1. When $k>2$, the expansion corresponding to the right of (4.1) is no longer convergent, but can be summed by RIESZ's means. WALFISZ has proved that the corresponding series in k-dimensions, namely

$$\sum\frac{r_k(n)J_{k/2}(2\pi\sqrt{nx})}{n^{k/4}} \tag{4.2}$$

is summable (n,δ) for $\delta>\frac{k-3}{2}$, and *not* summable for $\delta=\frac{k-3}{2}$. More complete results of this type were obtained by DIXON and FERRAR [7] and in a recent paper we obtained the following result [4, p. 248]: if

$$r_k(n,h)=\sum_{n_1^2+\dots+n_k^2=n}e^{2\pi i(n_1h_1+\dots+n_kh_k)},$$

then

$$\sum r_k(n,h)J_\mu(2\pi\xi\sqrt{n})\,n^l \tag{4.3}$$

is summable (n,η) for $\eta\geqq 0$ and $l<\frac{3}{4}-\frac{k}{2}+\frac{\eta}{2}$, $\mu>-1$ whenever ξ^2 is non-integral. (4.3) not only yields WALFISZ's result when $h_1=\dots=h_k=0$, but is actually sharper since μ does not depend on k. We will now show that a result which includes WALFISZ's and HARDY's, can be deduced as a direct consequence of theorem 1.

Corollary to theorem 1. *If ξ^2 is non-integral,*

$$\sum r_k(n)J_{k/2+\beta}(2\pi\xi\sqrt{n})\,n^l \tag{4.4}$$

is summable (n,η) for $\eta\geqq 0$ and $l<\frac{3}{4}-\frac{k}{2}+\frac{\eta}{2}$, $\beta>-1$.

Proof. It is known that the series

$$A+B\sum\frac{r_k(n)J_{k/2+\beta}(2\pi\xi\sqrt{n})}{n^{k/4+\beta/2}} \tag{4.5}$$

for suitable constants A and B, is the (spherical) multiple Fourier series of the function

$$f(x_1,\dots,x_k)=\sum\,[\xi^2-\{(n_1^2+x_1^2)+\dots+(n_k^2+x_k^2)\}]^\beta \tag{4.6}$$

at the origin $x=(0,\dots,0)$, for $\beta>-1$. [4, p. 243 (3.4)]. If ξ^2 is non-integral, the function given in (4.6) is infinitely differentiable in a neighborhood of the origin, and the terms of its Fourier series (4.5) satisfy the condition (in our notation of § 2)

$$A_n = O\left(n^{\frac{k-2}{2}+\varepsilon-\frac{k}{4}-\frac{\beta}{2}-\frac{1}{4}}\right) = O\left(n^{\frac{k-5}{4}-\frac{\beta}{2}+\varepsilon}\right) \tag{4.7}$$

since $J_\mu(x) = O(x^{-1/2})$ as $x \to \infty$, $\mu > -1$ and $r_k(n) = O\left(n^{\frac{k-2}{2}+\varepsilon}\right)$. Hence we apply theorem 1, and deduce that

$$\sum \frac{r_k(n)\, J_{k/2+\beta}(2\pi\xi\sqrt{n})}{n^{k/4+\beta/2}}\, n^p$$

is summable (n, η) for $\eta > \dfrac{k-3}{2} - \beta + 2p$. Setting $l = p - \dfrac{k}{4} - \dfrac{\beta}{2}$ we obtain that

$$\sum r_k(n)\, J_{k/2+\beta}(2\pi\xi\sqrt{n})\, n^l$$

is summable (n, η) for $\eta > 2l + k - \dfrac{3}{2}$ which is the required result.

Remark. The corollary will still hold when the order of the Bessel function in (4.4) is not necessarily $\dfrac{k}{2} + \beta$ but any $\mu > -1$; in order to see that, we have only to refer to the reasoning given in our previous paper [4, p. 246], which closely follows that of DIXON and FERRAR [7].

5. An improvement on theorem 1.

Theorem 1 was concerned with the case when the function $f(x)$ was infinitely differentiable in a neighborhood of a given point; we shall now prove a similar result in the case where the function has partial derivatives upto an assigned order which is finite; if this order exceeds the number $\frac{1}{4}(k-1)$, (where k is the dimension-number) then we already can reach the conclusion of theorem 1, without having to assume infinite differentiability. We shall however have a restriction on $\{a_{n_1 \ldots n_k}\}$ instead of on A_n. For the proof of the theorem we need the following

Lemma 5. *For given $\varepsilon > 0$ and $\varepsilon \leqq x \leqq 2\varepsilon$, let $\psi(x)$ be a function defined in the following way:*

(i) $\psi(\varepsilon) = 1$, $\psi(2\varepsilon) = 0$;

(ii) *$\psi(x)$ possesses derivatives of all orders in $\varepsilon \leqq x \leqq 2\varepsilon$;*

(iii) $\left(\dfrac{d^r\psi}{dx^r}\right)_{x=\varepsilon} = 0$, $\left(\dfrac{d^r\psi}{dx^r}\right)_{x=2\varepsilon} = 0$, *for* $r = 1, 2, 3, \ldots$.

Let $g(y)$ be defined in the following way:

(iv) $g(y) = 1$ *for* $|y| \leqq \varepsilon$;

(v) $g(y) = \psi(y)$ *for* $\varepsilon \leqq |y| \leqq 2\varepsilon$;

(vi) $g(y) = 0$ *for* $2\varepsilon \leqq |x| \leqq \pi$;

(vii) $g(y + 2\pi) = g(y)$.

Let $g(x_1, \ldots, x_k) = \prod_{r=1}^{k} g(x_r)$ *and let the Fourier expansion of* $g(x_1, \ldots, x_k)$ *be*

$$g(x_1, \ldots, x_k) \sim \sum_{-\infty}^{\infty} \cdots \sum b_{n_1 \ldots n_k} e^{i(n_1 x_1 + \ldots + n_k x_k)}.$$

Then for every $\beta > 0$ *we have:*

$$b_{n_1 \ldots n_k} = O\left(\frac{1}{(n_1^2 + \ldots + n_k^2)^\beta}\right). \tag{5.1}$$

Now let $f(x_1, \ldots, x_k)$ *be any periodic function having the period* 2π *in each variable and Lebesgue integrable, and let*

$$a_{n_1 \ldots n_k} = O\left(\frac{1}{(n_1^2 + \ldots + n_k^2)^\alpha}\right) \tag{5.2}$$

where $\{a_{n_1 \ldots n_k}\}$ *are the Fourier coefficients of* f. *If* $\{c_{n_1 \ldots n_k}\}$ *are the Fourier coefficients of the product* $f . g$, *then*

$$c_{n_1 \ldots n_k} = O\left\{\frac{1}{(n_1^2 + \ldots + n_k^2)^\alpha}\right\}. \tag{5.3}$$

Proof. An explicit example of a function ψ satisfying our requirements is found in WIENER [14, p. 562], where the interval $(0, 1)$ is considered instead of $(\varepsilon, 2\varepsilon)$.

Since $g(x_r)$ is infinitely differentiable, it follows by a well-known result in Fourier series that its Fourier coefficient

$$a_n^{(r)} = O\left(\frac{1}{n^{\beta_1}}\right)$$

for every $\beta_1 > 0$; from this it follows that (5. 1) is satisfied for every $\beta > 0$. Further, we have

$$c_{n_1 \ldots n_k} = \sum_{(m)=-\infty}^{\infty} \cdots \sum b_{m_1 \ldots m_k} a_{n_1 - m_1, \ldots, n_k - m_k}.$$

Hence

$$|c_{n_1 \ldots n_k}| = O\left(\sum \cdots \sum [\{1 + (n_1 - m_1)^2 + \ldots + (n_k - m_k)^2\}^{-\alpha} (1 + m_1^2 + \ldots + m_k^2)^{-\beta}]\right)$$

$$= O\left(\int_{-\infty}^{\infty} \cdots \int \{1 + (n_1 - \xi_1)^2 + \ldots + (n_k - \xi_k)^2\}^{-\alpha} \{1 + \xi_1^2 + \ldots + \xi_k^2\}^{-\beta} d\xi_1 \ldots d\xi_k\right).$$

If we subject the above integrand to an orthogonal transformation

$$\eta_r = \sum_s d_{rs} \xi_s$$

with determinant $+1$, and

$$d_{11} : d_{12} : \ldots : d_{1k} = n_1 : n_2 : \ldots : n_k,$$

then we have

$$\Sigma \xi_r^2 = \Sigma \eta_r^2, \; \Sigma n_r \xi_r = \sqrt{\Sigma n_r^2}\, \eta_1.$$

Hence

$$|c_{n_1 \dots n_k}| = O\Big(\int \dots \int [\{1 + (\sqrt{\Sigma n_r^2} - \eta_1)^2 + \dots + \eta_k^2\}^{-\alpha} \{1 + \eta_1^2 + \dots + \eta_k^2\}^{-\beta}]\, d\eta_1 \dots d\eta_k\Big)$$
$$= O\Big(\int \dots \int \{1 + (x - \eta_1)^2 + \dots + \eta_k^2\}^{-\alpha} \{1 + \eta_1^2 + \dots + \eta_k^2\}^{-\beta}\, d\eta_1 \dots d\eta_k\Big)$$

where $x = \sqrt{\Sigma n_r^2}$. Setting $|\eta| = \sqrt{\Sigma \eta_r^2}$, we have

$$|c_{n_1 \dots n_k}| = O\Big(\int_{|\eta| \leqq x/2} \!\!\dots\!\int + \int_{|\eta| > x/2} \!\!\dots\!\int\Big)$$

where

$$\int_{|\eta| \leqq x/2} \!\!\dots\!\int = O\left(\left\{1 + \left(\frac{x}{2}\right)^2\right\}^{-\alpha}\right) \int \dots \int \frac{d\eta_1 \dots d\eta_k}{(1 + \Sigma \eta_r^2)^\beta} = O(x^{-2\alpha}) = O[(\Sigma n_r^2)^{-\alpha}],$$

since β may be assumed large. Again,

$$\int_{|\eta| > x/2} \!\!\dots\!\int = O\Big(\int_{|\eta| > x/2} \!\!\dots\!\int \{1 + \Sigma \eta_r^2\}^{-\beta} d\eta_1 \dots d\eta_k\Big) =$$
$$= O\Big(\int_{t > x/2} t^{k-1} (1 + t^2)^{-\beta} dt\Big) = O(x^{k-2\beta}) = O(x^{-2\alpha}),$$

if we choose $\beta = \alpha + \dfrac{k}{2}$, and hence the lemma.

Theorem 2. *If $f(x)$ which is defined as in § 2 has continuous derivatives of total order $\leqq 2q$, where q is a non-negative integer, at a point x, and if*

$$a_{n_1 \dots n_k} = O((n_1^2 + \dots + n_k^2)^\beta),$$

then at that point, the series ΣA_n is summable (n, δ), $\delta \geqq 0$ and $\delta = \max(\eta, \gamma)$ where

$$\eta > \frac{\left(\frac{k-1}{2}\right)\left(\beta + \frac{k}{2}\right) - q}{\beta + \frac{k}{2} + q}$$

and $\gamma > 2\beta + k - 1$; in particular, if $q \geqq \dfrac{k-1}{4}$, then it is summable (n, δ) for any $\delta > 2\beta + k - 1$, $\delta \geqq 0$.

Proof. Without loss of generality we can assume the point in question to be the origin. We write the function $f(x)$ as follows:

$$f(x_1, \dots, x_k) = f(x_1, \dots, x_k)\, g(x_1, \dots, x_k) + [1 - g(x_1, \dots, x_k)]\, f(x_1, \dots, x_k)$$
$$= \varphi_1(x_1, \dots, x_k) + \varphi_2(x_1, \dots, x_k), \text{ say,}$$

where $g(x_1, \dots, x_k)$ is defined as in lemma 5. It follows from that lemma that $\varphi_2(x)$ is infinitely differentiable in a neighborhood of the origin (since it vanishes there), while $\varphi_1(x)$ is continuously differentiable $2q$ times everywhere. If we now write

$$\varphi_1 \sim \sum c_{n_1\ldots n_k} e^{i(n_1x_1+\ldots+n_kx_k)}$$

and

$$\varphi_2 \sim \sum d_{n_1\ldots n_k} e^{i(n_1x_1+\ldots+n_kx_k)}$$

by the same lemma 5, we have

$$\left.\begin{matrix} c_{n_1\ldots n_k} \\ d_{n_1\ldots n_k} \end{matrix}\right\} = O\left((n_1^2+\ldots+n_k^2)^\beta\right).$$

Theorem 1 is now applicable to φ_2, and so it follows that *its* Fourier expansion (summed spherically) is summable (n, γ) for

$$\gamma > 2\left(\beta+\frac{k-2}{2}\right)+1. \tag{5.4}$$

For φ_1 we proceed as follows. If we write

$$C_n = \sum_{n_1^2+\ldots+n_k^2=n} c_{n_1\ldots n_k} e^{i(n_1x_1+\ldots+n_kx_k)}$$

and apply the Laplace-operator q times, then owing to the continuity of the derivatives, it follows that $\Sigma C_n n^q$ is summable (n, δ) for $\delta > \frac{k-1}{2}$, at the origin. [2, Th. VI.] Hence it follows, as in the proof of theorem 1, that ΣC_n is summable (n, η) for

$$\eta > \delta - \frac{q(\delta+1)}{\left(\beta+\frac{k-2}{2}\right)+q+1}$$

or,

$$\eta > \frac{\left(\frac{k-1}{2}\right)\left(\beta+\frac{k}{2}\right)-q}{\beta+\frac{k}{2}+q}. \tag{5.5}$$

The first part of our theorem results from (5.4) and (5.5). If $2\beta+k-1<0$ and $\dfrac{\left(\frac{k-1}{2}\right)\left(\beta+\frac{k}{2}\right)-q}{\beta+\frac{k}{2}+q}<0$, then $\eta=\gamma=0$ so that $\delta=0$.

In order to prove the second part we note that summability (n, δ) of ΣA_n for *some* $\delta > 2\beta+k-1$ could only fail if

$$\frac{\left(\frac{k-1}{2}\right)\left(\beta+\frac{k}{2}\right)-q}{\beta+\frac{k}{2}+q} > 2\beta+k-1. \tag{5.6}$$

If we set

$$\beta+\frac{k-2}{2}=\alpha, \quad \frac{k-1}{2}=r \tag{5.7}$$

we see that (5. 6) is equivalent to

(5. 8) $$r(\alpha+1)>(\alpha+1)(2\alpha+1+2q).$$

Let us now discuss the following cases separately: (i) $\alpha+1=0$, (ii) $\alpha+1<0$, (iii) $0<\alpha+1<\frac{1}{2}$, (iv) $\alpha+1\geqq\frac{1}{2}$:

If $\alpha+1=0$ then the strict inequality in (5. 8) is impossible, and hence our theorem is proved in this case. If $\alpha+1<0$ then $-\beta>\frac{k}{2}$ so that ΣA_n converges absolutely, and our theorem is true trivially in this case. If $0<\alpha+1<\frac{1}{2}$ then we have $2\beta+k-1=2\alpha+1<0$; and since $\alpha+1+q>0$, and $r(\alpha+1)<q$ if $q\geqq\frac{k-1}{4}$, we also have

$$\frac{\left(\frac{k-1}{2}\right)\left(\beta+\frac{k}{2}\right)-q}{\beta+\frac{k}{2}+q}=\frac{r(\alpha+1)-q}{\alpha+1+q}<0.$$

Hence in this case $\eta=\gamma=0$, provided that $q\geqq\frac{k-1}{4}$, and we have convergence of ΣA_n, so the theorem is true. Finally if $\alpha+1\geqq\frac{1}{2}$ then $2\alpha+1\geqq0$: and if $q\geqq\frac{k-1}{4}$ then $r\leqq 2q$; so that we have $(2\alpha+1+2q)\geqq\frac{k-1}{2}$ which contradicts (5. 8); hence in this case also the theorem is proved.

6. Another convergence theorem.

We shall now establish a theorem more general than that of CHANDRASEKHARAN and MINAKSHISUNDARAM [6, Th. 4. 1]. We need the following lemmas.

Lemma 6. *Let $W(x)$ be a positive non-decreasing function of x, and $V(x)$ any positive function of x, both defined for $x>0$. Let $A(t)$ be a function of bounded variation in every finite interval, and*

$$A^k(t)=k\int_0^t (t-u)^{k-1}A(u)\,du,\ \ k>0.$$

Then

$$A(x+t)-A(x)=O[t^{\gamma}V(x)];\quad \gamma>0,\ t>0$$

and

$$A^k(x)=o(W)$$

together imply

$$A(x)=o\left(V^{\frac{k}{k+\gamma}}W^{\frac{\gamma}{k+\gamma}}\right).$$

If further, $V^{\frac{k}{k+\gamma}} W^{\frac{\gamma}{k+\gamma}}$ *is nondecreasing, then*

$$A^r(x) = o\left(V^{\frac{k-r}{k+\gamma}} W^{\frac{\gamma+r}{k+\gamma}}\right)$$

for $0 \leqq r \leqq k$.

This is a consequence of a convexity theorem of M. Riesz and the proof follows on well known lines. See [6, lemma 4. 2].

Lemma 7. (i) *If* $\delta > \frac{k-1}{2} + q$, *we have*

$$D_x^q T_R^\delta(x) = cR \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x+t)\, D_t^q [V_{k/2+\delta}(|t|\, R)]\, dt_1 \dots dt_k \tag{6. 1}$$

where $|t| = (t_1^2 + \dots + t_k^2)^{1/2}$, $V_\delta(x)$ *stands for* $J_\delta(x)/x^\delta$, J_δ *stands for the Bessel function of order* δ, *and* D_x^q, $f(x)$ *have the same meaning as in lemma* 2;

(ii) *if* $\delta > \frac{k-1}{2} + q$, *then*

$$R \int_{-\infty}^{\infty} \dots \int |f(x+t)\, D_t^q V_{k/2+\delta}(|t|\, R)|\, dt_1 \dots dt_k \leqq \tag{6. 2}$$

$$\leqq c \left[R^{k+q} \int_0^{1/R} |f(x, t)|\, t^{k-1} dt + R^{q+\frac{k}{2}-\delta-\frac{1}{2}} \int_{1/R}^{\infty} \frac{|f_x(t)|\, dt}{t^{\delta-\frac{k}{2}+\frac{3}{2}}} \right];$$

(iii) *if*

$$F(x, t) = \int_0^t |f(x, s)|\, s^{k-1-q} ds = o\,(t^{k+\theta}), \quad \theta > 0, \tag{6. 3}$$

as $t \to 0$, *then*

$$D_x^q T_R^\delta(x) = o\left(\frac{1}{R^\theta}\right)$$

as $R \to \infty$, *provided that* $\delta > \frac{k-1}{2} + q + \theta$.

Proof. Parts (i) and (ii) are contained in Bochner's paper [2, lemma 6, p. 349]; the argument for part (iii) runs parallel to Chandrasekharan's [5, Th. V]. We have only to consider the right side of (6. 2). Assumption (6. 3) yields

$$R^{k+q} \int_0^{1/R} |f(x, s)|\, s^{k-1} ds = o\,(R^{-\theta}), \tag{6. 4}$$

and as for the second integral, we split it in two; setting $\varrho = \delta - q - \frac{k}{2} + \frac{1}{2}$, we have

$$R^{-\varrho} \left[\int_{1/R}^{\eta} + \int_{\eta}^{\infty} \right] = \varphi_1 + \varphi_2, \text{ say.}$$

To estimate φ_2 we have only to use the fact that

$$G(x, u) = \int_0^u t^{k-1} |f(x, t)| \, dt = O(u^k)$$

as $u \to \infty$. See [5, p. 213 (2. 11)]. For,

$$\varphi_2 = \frac{1}{R^\varrho} \int_\eta^\infty \frac{t^{k-1} |f(x, t)| \, dt}{t^{k+q+\varrho}} = \frac{1}{R^\varrho} \left[\left\{ \frac{G(x, t)}{t^{k+q+\varrho}} \right\}_\eta^\infty + c \int_\eta^\infty \frac{G(x, t) \, dt}{t^{k+q+\varrho+1}} \right] = O\left(\frac{1}{R^\varrho}\right) = o\left(\frac{1}{R^\theta}\right)$$

provided that

$$\varrho > \theta, \quad \text{or} \quad \delta > \frac{k-1}{2} + q + \theta. \tag{6. 5}$$

As for φ_1,

$$\varphi_1 = R^{-\varrho} \int_{1/R}^\eta t^{-k-\varrho} \, dF(t),$$

and we now integrate by parts, and use (6. 3) in the same manner as in [5, p. 219 (3. 26)], thus obtaining

$$\varphi_1 = o(R^{-\theta}) \quad \text{if} \quad \varrho > \theta.$$

This concludes the proof of the lemma.

Theorem 3. *If $f(x)$ is defined as in § 2, and if at a given point x,*

$$\frac{1}{t^k} \int_0^t |f(x, s)| \, s^{k-1-2q} \, ds = o(t^\theta), \quad \theta \geqq 0 \tag{6. 6}$$

as $t \to 0$, where q is a non-negetive integer, and if

$$A_n = O(n^\alpha) \tag{6. 7}$$

then $\Sigma A_n n^q$ is summable (n, r) for $\delta \geqq r \geqq 0$ provided that θ, α, q and r satisfy the relation

$$2(\delta - r)(\alpha + q + 1) - \theta(1 + r) = 0 \tag{6. 8}$$

for some $\delta > \dfrac{k-1}{2} + 2q + \theta$.

Proof. By lemma 7 (iii), assumption (6. 6) implies that

$$\Delta_x^q T_R^\delta(x) = o(R^{-\theta})$$

or

$$\Delta_x^q S_R^\delta(x) = o(R^{2\delta - \theta}) \tag{6. 9}$$

where $\delta > \dfrac{k-1}{2} + 2q + \theta$ and $\Delta_x^q = \left(\dfrac{\partial^2}{\partial x_1^2} + \ldots + \dfrac{\partial^2}{\partial x_k^2} \right)^q$.

If we now set

$$B_n = A_n n^q, \qquad B(R) = \sum_{n<R^2} B_n, \qquad B^\delta(R) = 2\delta \int_0^R (R^2 - u^2)^{\delta-1} B(u)\, u\, du, \ \delta > 0,$$

as in § 2, then (6. 9) implies

(6. 10) $$B^\delta(\sqrt{R}) = o(R^{\delta-\theta/2})$$

for $\delta > \frac{k-1}{2} + 2q + \theta$. On the other hand, $B_n = O(n^{\alpha+q})$ and hence

(6. 11) $$|B(\sqrt{\omega+t}) - B(\sqrt{\omega})| \leqq \sum_{\omega\leqq n\leqq\omega+t} B_n = O\Big(\sum_{\omega\leqq n\leqq\omega+t} n^{\alpha+q}\Big) = O(t\omega^{\alpha+q}).$$

From (6. 10) and (6. 11) it follows that we can apply lemma 6 if we choose $B(\sqrt{x}) = A(x)$, $x^{\alpha+q} = V(x)$ and $x^{\delta-\theta/2} = W(x)$, and then we obtain

(6. 12) $$B^r(\sqrt{R}) = o(R^\beta)$$

where $0 \leqq r \leqq \delta$ and

(6. 13) $$\beta = \frac{(\alpha+q)(\delta-r)}{\delta+1} + \frac{(1+r)(\delta-\theta/2)}{\delta+1}.$$

Let us write (6. 12) in the form

(6. 14) $$\frac{B^r(R)}{R^{2r}} = o(R^{2\beta-2r}) = o(R^\eta), \text{ say.}$$

If $\eta = 0$, then it follows that $\Sigma A_n n^q$ is summable (n, r); this will be the case if

(6. 15) $$2(\delta-r)(\alpha+q+1) - \theta(1+r) = 0$$

where $\delta > \frac{k-1}{2} + 2q + \theta$.

Remarks. (1) Let us write relation (6. 15) in the form

(6. 16) $$r = \frac{2\delta(\alpha+q+1)-\theta}{\theta+2(\alpha+q+1)}.$$

Now if $\theta = 0$ then $r = \delta$, where $\delta > \frac{k-1}{2} + 2q$. Thus we obtain BOCHNER's result [2, Th. I] as a special case.

(2) Let $q = 0$ and $k = 2$. Then it follows from (6. 16) that $r = 0$, if $2\delta(\alpha+1) = \theta$ where $\delta > \theta + \frac{1}{2}$; and this will be the case if $\alpha < \frac{\theta}{2\theta+1} - 1$.

Suppose now that

(6. 17) $$a_{n_1 n_2} = O\left(\frac{1}{(n_1^2+n_2^2)^p}\right);$$

then $A_n = O(n^{\varepsilon-p})$ for every $\varepsilon > 0$, since $r_k(n) = O\left(n^{\frac{k-2}{2}+\varepsilon}\right)$. Hence under the assumption (6. 6) with $q = 0$, $k = 2$ and the assumption (6. 17) we conclude that ΣA_n converges if $\varepsilon - p < \frac{\theta}{2\theta+1} - 1$ for every $\varepsilon > 0$ or if

(6. 18) $$p > 1 - \frac{\theta}{2\theta+1},$$

which is exactly the theorem of CHANDRASEKHARAN and MINAKSHISUNDARAM [6, Th. 4. 1].

(3) Though in the assumption (6. 6) we have q as an integer, we can, if necessary, determine the order of summability of $\Sigma A_n n^h$ for *arbitrary* h by applying ANANDA-RAU's theorem (lemma 1). We choose not to repeat this kind of computation.

(4) Our hypothesis (6. 6) differs from the hypothesis in theorems 1 and 2 in as much as it governs the behaviour of the function $f(x)$ at a given point x, and not in a whole neighborhood of it.

References.

1. K. ANANDA-RAU, On the convergence and summability of Dirichlet's series, *Proceedings London Math. Soc.*, (2) **34** (1932), pp. 414—440.
2. S. BOCHNER, Summation of multiple Fourier series by spherical means, *Transactions American Math. Soc.*, **40** (1936), pp. 175—207.
3. S. BOCHNER, Summation of derived Fourier series, an application to Fourier expansions on compact Lie groups, *Annals of Math.*, **37** (1936), pp. 345—356.
4. S. BOCHNER and K. CHANDRASEKHARAN, Summations over lattice points in k-space, *Quarterly Journal of Math.*, **19** (1948), pp. 238—248.
5. K. CHANDRASEKHARAN, On the summation of multiple Fourier series. I, *Proceedings London Math. Soc.*, (2) **50** (1948), pp. 210—222.
6. K. CHANDRASEKHARAN and S. MINAKSHISUNDARAM, Some results on double Fourier series, *Duke Math. Journal*, **14** (1947), pp. 731—753.
7. A. L. DIXON and W. L. FERRAR, Some summations over lattice points of a circle. (I) and (II), *Quarterly Journal of Math.*, **5** (1934), pp. 48—63, 172—185.
8. G. H. HARDY, The average order of arithmetical functions $P(x)$ and $\Delta(x)$, *Proceedings London Math. Soc.*, (2) **15** (1916), pp. 192—213.
9. G. H. HARDY and E. LANDAU, The lattice points of a circle, *Proceedings Royal Soc.*, (A) **105** (1924), pp. 244—258.
10. G. H. HARDY and J. E. LITTLEWOOD, Some new convergence criteria for Fourier series, *Annali della R. Scuola Normale Superiore di Pisa*, (2) **3** (1934), pp. 43—62.
11. A. OPPENHEIM, Somme identities in the theory of numbers, *Proceedings London Math. Soc.*, (2) **26** (1927), pp. 295—350.
12. A. WALFISZ, Über summatorische Funktionen einiger Dirichletscher Reihen, *Dissertation Göttingen* (1922).
13. A. WALFISZ, Über das Piltzsche Teilerproblem in algebraischen Zahlkörpern, *Math. Zeitschrift*, **22** (1925), pp. 153—188.
14. N. WIENER, The operational calculus, *Math. Annalen*, **95** (1926), pp. 557—584.
15. J. R. WILTON, The lattice points of a circle. An historical account of the problem, *Messenger of Math.*, **58** (1929), pp. 67—80.
16. J. R. WILTON, A series of Bessel functions connected with the theory of lattice points, *Proceedings London Math. Soc.*, (2) **29** (1929), pp. 168—188.

PRINCETON UNIVERSITY,
TATA INSTITUTE OF FUNDAMENTAL RESEARCH.

(*Received July 12, 1949.*)

On the discreteness of the spectrum of a differential equation.

By E. C. TITCHMARSH in Oxford.

It was proved by WEYL[1]) that the spectrum associated with the differential equation

(1) $$\frac{d^2\varphi}{dx^2}+\{\lambda-q(x)\}\varphi=0 \qquad (0\leqq x<\infty)$$

is discrete if $q(x)$ is bounded in any finite interval and tends to infinity as $x\to\infty$. His proof is reproduced in my book *Eigenfunction Expansions Associated with Second Order Differential Equations* (Oxford, 1946), § 5. 12. Other proofs have since been given[2]).

The following is another simple proof. Let $\varphi(x,\lambda)$ be the solution of (1) which satisfies a given boundary condition

(2) $$\varphi(0,\lambda)\cos\alpha+\varphi_x(0,\lambda)\sin\alpha=0$$

at $x=0$. Then it can be proved as in § 5. 12 of my book that, for every real λ, either $\varphi(x,\lambda)$ is L^2 (in which case $\varphi(x,\lambda)$ and $\varphi_x(x,\lambda)$ both tend to zero as $x\to\infty$), or $\varphi(x,\lambda)\to\infty$, or $\varphi(x,\lambda)\to-\infty$. Consider any finite interval $a\leqq\lambda\leqq b$, and denote the sub-sets of this interval where φ has the above properties by E_0, E_1, and E_2 respectively.

If λ' belongs to E_1, $\varphi(x,\lambda')\to\infty$, $\varphi_{xx}(x,\lambda')\to\infty$ (by (1)), and so $\varphi_x(x,\lambda')\to\infty$. Hence for some ξ, with $q(\xi)>\lambda'$, $\varphi(\xi,\lambda')>0$ and $\varphi_x(\xi,\lambda')>0$. Hence also $\varphi(\xi,\lambda)>0$ and $\varphi_x(\xi,\lambda)>0$ if $\lambda-\lambda'$ is sufficiently small. This, however, implies that $\varphi(x,\lambda)\to\infty$. Hence any point of E_1 is an interior point of an interval of points of E_1, and so E_1 is an open set. Similarly E_2 is an open set. Hence E_0 is a closed set.

---

[1]) H. WEYL, Über gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkürlicher Funktionen, *Math. Annalen*, **68** (1910), pp. 220—269.

[2]) K. FRIEDRICHS, Spektraltheorie halbbeschränkter Operatoren und Anwendung auf die Spektralzerlegung von Differentialoperatoren. II., *Math. Annalen*, **109** (1934), pp. 685 - 713. Criteria for the discrete character of the spectra of ordinary differential operators, *Studies and Essays presented to R. Courant* (New York, 1948), pp. 145—160; a proof by E. C. TITCHMARSH will be published in the *Annali di Matematica*.

The above argument also shows that, if λ' is a point of E_1, then, for some ξ, $\varphi(\xi, \lambda) \geqq m > 0$, while $\varphi(x, \lambda)$ is steadily increasing for $x \geqq \xi$, if λ is in some interval $|\lambda - \lambda'| \leqq \eta$. Hence $\varphi(x, \lambda) \geqq m$ for $x \geqq \xi$, $|\lambda - \lambda'| \leqq \eta$. It follows as on p. 116 of my book that the function $k(\lambda)$ is constant throughout the interval $|\lambda - \lambda'| < \eta$, and so in fact is constant throughout each interval of E_1.

To prove the theorem, we have now to show that E_0 consists at most of a finite number of points.

Suppose on the contrary that there is a sequence of values of λ tending to (but different from) a limit μ, such that these λ's and μ all belong to E_0. Let x_1 be such that $q(x) \geqq \mu + \delta$ $(\delta > 0)$ for $x \geqq x_1$ (such an x_1 exists if $q(x) \to \infty$). Let $x_2 > x_1$ be such that $\varphi(x_2, \mu) \neq 0$, and suppose e. g. that $\varphi(x_2, \mu) > 0$. As in § 5.12 of my book, this implies (since $\varphi(x, \mu)$ is L^2) that $\varphi(x, \mu)$ decreases steadily to zero for $x \geqq x_2$, and in particular that $\varphi(x, \mu) > 0$ for $x \geqq x_2$.

Now $\varphi(x, \lambda)$ is a continuous function of both variables in any finite region (cf. § 1.5 of my book), and so $\varphi(x, \lambda) \to \varphi(x, \mu)$ as $\lambda \to \mu$, uniformly over $0 \leqq x \leqq x_2$. Hence

$$\lim_{\lambda \to \mu} \int_0^{x_2} \varphi(x, \lambda)\, \varphi(x, \mu)\, dx = \int_0^{x_2} \{\varphi(x, \mu)\}^2\, dx > 0,$$

and so

$$\int_0^{x_2} \varphi(x, \lambda)\, \varphi(x, \mu)\, dx > 0$$

if λ is sufficiently near to μ.

Also

$$\varphi(x_2, \lambda) \to \varphi(x_2, \mu) > 0$$

and so $\varphi(x_2, \lambda) > 0$ if λ is sufficiently near to μ. Since $q(x) - \lambda > 0$ if $x \geqq x_1$ and λ is sufficiently near to μ, this implies, as before, that $\varphi(x, \lambda) > 0$ for $x \geqq x_2$ and λ sufficiently near to μ. Hence

$$\int_{x_2}^{\infty} \varphi(x, \lambda)\, \varphi(x, \mu)\, dx > 0$$

if λ is sufficiently near to μ. Altogether

$$\int_0^{\infty} \varphi(x, \lambda)\, \varphi(x, \mu)\, dx > 0 \tag{3}$$

if λ is sufficiently near to μ.

This, however, is impossible; for on multiplying (1), and the corresponding equation with μ, by $\varphi(x, \mu)$, $\varphi(x, \lambda)$, respectively and subtracting, we obtain

$$(\lambda-\mu)\,\varphi(x,\lambda)\,\varphi(x,\mu)=\frac{\partial}{\partial x}\{\varphi(x,\lambda)\,\varphi_x(x,\mu)-\varphi(x,\mu)\,\varphi_x(x,\lambda)\}.$$

Using (2) it follows that

$$(\lambda-\mu)\int_0^X \varphi(x,\lambda)\,\varphi(x,\mu)\,dx=\varphi(X,\lambda)\,\varphi_x(X,\mu)-\varphi(X,\mu)\,\varphi_x(X,\lambda),$$

which tends to 0 as $X\to\infty$. Since $\lambda\neq\mu$ it follows that

(4) $$\int_0^\infty \varphi(x,\lambda)\,\varphi(x,\mu)\,dx=0.$$

This gives a contradiction, and the theorem follows.

(Received July 9, 1949.)

On the differentiability of semi-group operators.

By EINAR HILLE in New Haven, Conn.

1. Let $\mathfrak{X}$ be a complex (B)-space, A a closed linear operator on $\mathfrak{X}$ to itself whose domain $\mathfrak{D}[A]$ is dense in $\mathfrak{X}$. Suppose that the operator $\lambda I - A$ has a bounded inverse $R(\lambda; A)$ for each fixed real positive value of λ and that

$$(1.1) \qquad \lambda \| R(\lambda; A) \| \leqq 1, \quad \lambda > 0.$$

Under these assumptions it is known (see E. HILLE [1], p. 238 and K. YOSIDA [2], p. 15) that A is the infinitesimal generator of a semi-group $\mathfrak{S} = \{T(\xi)\}$, $\xi > 0$, of linear bounded operators $T(\xi)$ with the properties

(i) $\quad T(\xi_1)\, T(\xi_2) = T(\xi_1 + \xi_2),\ \xi_1 > 0,\ \xi_2 > 0,$

(ii) $\quad \| T(\xi) \| \leqq 1,$

(iii) $\quad \lim\limits_{\xi \to 0} T(\xi) x = x, \quad x \in \mathfrak{X}.$

Here (iii) implies the further property $\lim\limits_{\xi \to \xi_0} T(\xi) x = T(\xi_0) x$ for $\xi_0 > 0$. Conversely, if a semi-group $\mathfrak{S}$ with the properties (i), (ii) and (iii) is given, then it has an infinitesimal generator A which is a linear closed operator whose domain $\mathfrak{D}[A]$ is dense in $\mathfrak{X}$. Further the resolvent $R(\lambda; A)$ exists for $\Re(\lambda) > 0$ and

$$(1.2) \qquad \sigma \| R(\sigma + i\tau; A) \| \leqq 1, \quad \lambda = \sigma + i\tau, \quad \sigma > 0.$$

For $x \in \mathfrak{D}[A]$ we have

$$(1.3) \qquad \lim_{\delta \to 0} \frac{1}{\delta} [T(\xi + \delta) x - T(\xi) x] = T'(\xi) x = A\, T(\xi) x = T(\xi) A x,$$

in the sense of strong convergence.

The assumptions on A stated above imply that $T(\xi)$ is strongly continuous for $\xi \geqq 0$ and no further continuity properties may be asserted in general. Similarly, for $\xi > 0$ the operator $T'(\xi) = A\, T(\xi)$ is ordinarily an unbounded operator whose domain of definition contains $\mathfrak{D}[A]$ and may coincide with $\mathfrak{D}[A]$. For the higher derivatives we have a similar situation; the domain of $T^{(n)}(\xi) = A^n\, T(\xi)$ contains $\mathfrak{D}[A^n]$ which is dense in $\mathfrak{X}$ for every n and $\bigcap_n \mathfrak{D}[A^n]$ is also dense in $\mathfrak{X}$.

Thus if we want to get semi-group operators with stronger continuity and differentiability properties, we must impose stronger restrictions on A. Two sets of such conditions were given in § 12. 2 of [1]. The first set give continuity of $T(\xi)$ in the uniform operator topology for $\xi > 0$, but not for $\xi = 0$, nor the existence of bounded derivatives, while the second set implies the existence of derivatives of all orders for $\xi > 0$ but not analyticity. In view of this situation we shall investigate the existence of derivatives of semi-group operators in the present note.

2. We start with

Theorem 1. *If for a positive ξ_0 the semi-group operator $T(\xi_0)$ maps $\mathfrak{X}$ upon a subset of $\mathfrak{D}[A]$, then $T'(\xi) = AT(\xi)$ exists as a bounded operator for $\xi \geqq \xi_0$. Moreover, $T^{(n)}(\xi) = A^n T(\xi)$ exists as a bounded operator for $\xi \geqq n\xi_0$, $n = 1, 2, 3, \ldots$.*

Proof. $AT(\xi_0)$ is a linear closed operator which is defined everywhere in $\mathfrak{X}$, hence it is bounded. Since $AT(\xi) = AT(\xi_0)\,T(\xi - \xi_0)$ for $\xi > \xi_0$, it follows that $AT(\xi)$ is also bounded. Further, for $\xi \geqq n\xi_0$ we have

$$(2.1) \qquad T^{(n)}(\xi) = A^n T(\xi) = \left[AT\left(\frac{\xi}{n}\right)\right]^n, \quad n = 1, 2, 3, \ldots$$

so that the higher derivatives exist as asserted.

Corollary. *If $T(\xi)[\mathfrak{X}] \subset \mathfrak{D}[A]$ for each $\xi > 0$, then $T^{(n)}(\xi)$ exists as a bounded operator on $\mathfrak{X}$ to $\mathfrak{X}$ for each $\xi > 0$ and $n = 1, 2, 3, \ldots$*

Conversely, if $T'(\xi)$ exists as a bounded operator for $\xi = \xi_0$, then the limit in the first member of (1.3) exists for all x when $\xi = \xi_0$, that is, $T(\xi_0)[\mathfrak{X}] \subset \mathfrak{D}[A]$ so that the condition of Theorem 1 is necessary as well as sufficient.

Theorem 2. *If $T(\xi)$ satisfies* (ii) *and if $T'(\xi)$ exists as a bounded operator for $\xi > \xi_0$, then $\|T'(\xi)\|$ is a monotone decreasing function of ξ in (ξ_0, ∞).*

For if $\delta > 0$ then $\|T'(\xi+\delta)\| = \|T'(\xi)\,T(\delta)\| \leqq \|T'(\xi)\|$. The same conclusion obviously holds for $\|T^{(n)}(\xi)\|$ in $(n\xi_0, \infty)$.

In particular, it follows that $\|T'(\xi)\|$ tends to a finite or infinite limit when ξ decreases to ξ_0. If $\xi_0 > 0$, it may very well happen that $\lim\limits_{\xi \to \xi_0} \|T'(\xi)\|$ is finite. In order to see this, we shall introduce a class of operators which will be used repeatedly in the following (cf. [1], Theorems 18.2.1 and 18.4.1). Let $\mathfrak{X} = L_2(-\infty, \infty)$, let $F(u)$ be the Fourier transform of $f(t) \in \mathfrak{X}$ and set

$$(2.2) \qquad T(\xi)[f] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\xi\varrho(u) + itu} F(u)\, du,$$

where $\varrho(u)$ is a continuous function whose real part is never negative and the integral exists in the sense of mean convergence. This defines a semi-group of linear operators in $\mathfrak{X}$ and $\|T(\xi)\|$ is the essential supremum of

$|e^{-\xi\varrho(u)}|$ and hence $\leqq 1$. Here we choose

$$\varrho(u) = |u| + ie^{\alpha|u|},\ \alpha > 0.$$

A simple calculation shows that $T'(\xi)$ exists as a bounded operator for $\xi \geqq \alpha$ but not for $\xi < \alpha$. The situation is different for $\xi_0 = 0$ since $T'(0)$ is bounded if and only if A is bounded in which case $T(\xi) = \exp(\xi A)$, that is, $T(\xi)$ is an entire function of ξ of exponential type.

We shall now consider the case in which $\xi_0 = 0$, that is, $T'(\xi)$ exists as a bounded operator for $\xi > 0$. The rate of growth of $\|T'(\xi)\| \equiv g(\xi)$ as $\xi \to 0$ is of fundamental importance for the following. We have observed above that $g(\xi)$ can stay bounded if and only if A is a bounded operator. Actually a sharper result holds.

Theorem 3. *A is bounded and $T(\xi) = \exp(\xi A)$ if*

$$\limsup_{\xi \to 0} \xi \|T'(\xi)\| < \frac{1}{e}. \tag{2.3}$$

Proof. Since $T(\xi)$ has derivatives of all orders, we may use Taylor's theorem obtaining

$$T(\xi) = \sum_{k=0}^{n-1} \frac{(\xi-\alpha)^k}{k!} A^k T(\alpha) + \frac{1}{(n-1)!} \int_{\alpha}^{\xi} (\xi-\eta)^{n-1} A^n T(\eta)\, d\eta. \tag{2.4}$$

Using (2.1) and (2.3) we see that the remainder tends to zero when $n \to \infty$ provided $0 < \alpha \leqq \xi < \alpha\left(1 + \frac{1}{\varrho}\right)$ where ϱ equals e times the left member of (2.3). It follows that the Taylor expansion converges for these values of ξ and represents $T(\xi)$. But the power series converges for $|\xi - \alpha| < \frac{\alpha}{\varrho}$ and in this circle it defines a semi-group operator which is the analytic continuation of $T(\xi)$. It follows that $T(\xi)$ is analytic in some neighborhood of the origin and this requires that $T(\xi)$ is an entire function so that $T(\xi) = \exp(\xi A)$ with a bounded operator A.

Corollary. *If $T(\xi)$ is a proper semi-group operator which has a bounded derivative for each positive ξ then*

$$\limsup_{\xi \to 0} \xi \|T'(\xi)\| \geqq \frac{1}{e}. \tag{2.5}$$

This is actually the best possible result of its kind, for if we take $\varrho(u) = |u|$ in (2.2), that is, if we form the Poisson-Abel transform of the Fourier integral, then $\xi\|T'(\xi)\| \equiv \frac{1}{e}$ for $\xi > 0$. On the other hand, there is no upper limit for the rate of growth of $\|T'(\xi)\|$ when $\xi \to 0$. In order to see this, we have merely to choose $\varrho(u) = |u| + ie^{\varphi(u)}$ where $\varphi(u) = o(u)$. The slower $u^{-1}\varphi(u) \to 0$ when $|u| \to \infty$, the faster grows the norm of $T'(\xi)$ when

$\xi \to 0$ and by a suitable choice of $\varphi(u)$ we can achieve that the norm grows faster than a preassigned function of ξ.

3. Conditions (i), (ii) and (iii) imply that $R(\lambda; A)$ exists for $\sigma > 0$ and satisfies (1. 2). Ordinarily no more may be asserted, but if $T(\xi)$ has derivatives, the resolvent set becomes more extensive.

Theorem 4. *Suppose that $T(\xi)$ is differentiable for $\xi > 0$ and set*

$$\| T'(\xi) \| = g(\xi) = \xi G(\xi). \tag{3. 1}$$

Let $\delta(\tau)$ be the distance of the point $1 + \tau i$ from the spectrum of A. For large values of $|\tau|$ we have

$$\delta(\tau) > \frac{1}{3\eta(\tau)}, \tag{3. 2}$$

where $\eta(\tau)$ is the unique root of the equation

$$G(\eta) = |\tau|. \tag{3. 3}$$

Proof. Since

$$R(\lambda; A) = \sum_{n=0}^{\infty} [R(\lambda_0; A)]^{n+1} (\lambda_0 - \lambda)^n$$

converges for $|\lambda - \lambda_0| \, \|R(\lambda_0; A)\| < 1$, it follows that

$$\delta(\tau_0) \, \|R(\lambda_0; A)\| \geqq 1, \quad \lambda_0 = 1 + i\tau_0.$$

Thus (3. 2) is implied by

$$\|R(1 + \tau i; A)\| < 3\eta(\tau). \tag{3. 4}$$

But

$$R(\lambda; A) = \int_0^{\infty} e^{-\lambda\xi} T(\xi) \, d\xi = \int_0^{\eta} + \int_{\eta}^{\infty} = J_1 + J_2,$$

where $\eta = \eta(\tau)$ is chosen as indicated above. We note that $G(\xi)$ is strictly decreasing from $+\infty$ to 0 when ξ goes from 0 to $+\infty$, so the equation (3. 3) has a unique root and if $|\tau|$ is sufficiently large, $\eta(\tau)$ is less than one. Without restricting the generality, we may assume that (2. 5) holds so that $\eta(\tau)$ is at least $O(|\tau|^{-\frac{1}{2}})$ when $|\tau| \to \infty$. For J_1 we have the trivial estimate $\|J_1\| \leqq \eta$. An integration by parts gives

$$\|J_2\| \leqq \frac{1}{|\lambda|} [1 + g(\eta)].$$

In view of the choice of $\eta(\tau)$ we see that (3. 4) holds for large values of $|\tau|$ and consequently also the desired relation (3. 2).

The resulting estimate is not particularly accurate, mainly on account of the crude estimate used for J_1. The latter integral is of the same order of magnitude as

$$\int_0^{\eta} \left[T\left(\xi + \frac{\pi}{|\tau|}\right) - T(\xi) \right] d\xi.$$

This suggests that a study of the modulus of continuity of $T(\xi)$ in $L(0, \eta)$ might lead to further improvements of the estimate. We shall not pursue this possibility here, however. The same method as used above leads to

Theorem 5. *If $T(\xi)$ is differentiable for $\xi > 0$ and if $\log g(\xi) \in L(0,1)$, then*

$$\left\{ \int_{-\infty}^{-1} + \int_{1}^{\infty} \right\} \|R(1+\tau i; A)\| \frac{d\tau}{|\tau|} < \infty. \tag{3.5}$$

Proof. If $\log g(\xi) \in L(0, 1)$ so does $\log G(\xi)$. Without restricting the generality we may suppose that $G(\xi)$ is absolutely continuous, since otherwise we may replace $G(\xi)$ by an absolutely continuous dominant having the same integrability properties. In view of (3.4) the integral in (3.5) is dominated by a constant multiple of

$$\int_{1}^{\infty} \eta(\tau) \frac{d\tau}{\tau} = -\int_{0}^{\eta_0} \eta \frac{G'(\eta)}{G(\eta)} d\eta = \int_{0}^{\eta_0} \log G(\eta)\, d\eta.$$

Here we have used the fact that $\eta \log G(\eta) \to 0$ with η and that $G(\eta_0) = 1$. This completes the proof.

The condition $\log g(\xi) \in L(0, 1)$ is probably far too restrictive for the desired conclusion. In fact there are transformations of type (2.2) for which (3.5) holds and merely $\log\log g(\xi) \in L(0, 1)$. For this class of transformations the condition $\log\log g(\xi) \in L(0, 1)$ is the best possible of its kind which will ensure convergence of (3.5). In order to prove an improved version of Theorem 5 with $\log g(\xi)$ replaced by $\log\log g(\xi)$ it would be sufficient to prove that

$$\|J_1\| \leqq \frac{\eta}{\log|\tau|}. \tag{3.6}$$

4. Theorem 4 suggests that if $\delta(\tau)$ grows sufficiently fast with $|\tau|$, then the semi-group operator $T(\xi)$ generated by A might be differentiable. We shall prove

Theorem 6. *If the operator A satisfies the conditions of section 1, if $\|R(\lambda; A)\| \to 0$ when $\lambda \to \infty$ in such a manner that the distance of λ from the spectrum of A becomes infinite, and if, for every fixed positive K, the inequality $\delta(\tau) > K \log|\tau|$ holds except in a set of intervals over which the total variation of τ^2 is finite, then the operator $T(\xi)$ generated by A has derivatives of all orders for $\xi > 0$.*

Proof. We know in advance that $T(\xi)$ exists and has the properties (i), (ii), and (iii). By Theorem 1 it is sufficient to prove the existence of $T'(\xi)$ for $\xi > 0$. For this purpose we consider the integral $\int e^{\lambda\xi} R(\lambda; A)\, d\lambda$ taken along a closed contour $PQRSP$ where PQ, QR, and RS are straight line segments and SP is an arc of the curve Γ: $\lambda = 1 - \frac{1}{2}\delta(\tau) + \tau i$, $Q = 1 - \omega i$

$R = 1 + \omega i$ while the imaginary part of P equals that of Q and the imaginary part of S equals that of R. We let $\omega \to \infty$; using the fact that $\|R(\lambda; A)\|$ tends uniformly to zero on and to the right of Γ, one sees that the integrals along the horizontal line segments PQ and RS tend to zero. That the integral along the arc PS of the curve Γ tends to a limit when $\omega \to \infty$ follows from the absolute convergence of the resulting integral which in its turn follows from the absolute convergence of the integral (4. 2) discussed below. It follows that the integral from Q to R along the vertical line tends to a limit in the uniform operator topology when $\omega \to \infty$. But for $x \in \mathfrak{D}[A]$ we have (see Theorem 11. 7. 1 of [1])

$$T(\xi)x = \lim_{\omega \to \infty} \frac{1}{2\pi i} \int_{1-\omega i}^{1+\omega i} e^{\lambda \xi} R(\lambda; A) x \, d\lambda,$$

whence it follows that

$$T(\xi) = \frac{1}{2\pi i} \int_{\Gamma} e^{\lambda \xi} R(\lambda; A) \, d\lambda. \tag{4.1}$$

Formally the derivative of $T(\xi)$ is given by

$$T'(\xi) = \frac{1}{2\pi i} \int_{\Gamma} e^{\lambda \xi} \lambda R(\lambda; A) \, d\lambda, \tag{4.2}$$

and all we have to do in order to prove the theorem is to show the absolute convergence of this integral for $\xi > 0$. The norm of this integral is dominated by a constant multiple of

$$e^{\xi} \int_{-\infty}^{\infty} |\tau| e^{-\frac{1}{2}\xi \delta(\tau)} d\tau. \tag{4.3}$$

The range of integration may now be split into two subsets E_1 and E_2. In E_1 we shall have $\delta(\tau) > (6/\xi) \log |\tau|$ and $|\tau| > 1$ so that the integral over E_1 converges as $\int_1^{\infty} \tau^{-2} d\tau$, while the integral over E_2 is dominated by the total variation of τ^2 over E_2 which is finite by assumption. This completes the proof.

References.

[1] E. Hille, Functional Analysis and Semi-groups. *American Math. Society Colloquium Publications,* Vol. XXXI, (New York, 1948).

[2] K. Yosida, On the differentiability and the representation of one-parameter semi-group of linear operators, *Journal Math. Society of Japan,* 1 (1948), pp. 15—21.

(Received August 5, 1949.)

Some prime-number consequences of the Ikehara theorem.

By NORBERT WIENER and LEONARD GELLERT in Cambridge, Mass.

SELBERG[1] and ERDŐS[2] have recently shown the Prime Number Theorem to be demonstrable by elementary methods[1]). The present paper is devoted, not to elementary proofs of this theorem, but to simple analytical considerations which may throw light on the reasons why such elementary, but not easy, proofs may be expected to function. Our fundamental tool is the Ikehara theorem[3]), to the effect that if for $\Re z > 1$,

$$(1) \qquad \int_0^\infty u^{-z} dF(z) = \frac{a}{z-1} + G(z),$$

where $F(z)$ is monotone; and if over every finite range of x,

$$(2) \qquad \lim_{x \to 1+0} G(x+iy) = H(y),$$

where $H(y) \in L$ over every finite range; then

$$(3) \qquad \lim_{A \to \infty} \frac{1}{A} F(A) = a.$$

We shall consider the two Dirichlet series:

$$(4) \qquad \int_0^\infty u^{-z} \log u \; d\tilde{\omega}(u) = -\frac{\zeta'(z)}{\zeta(z)},$$

and

$$-\int \frac{\zeta''(z)}{\zeta(z)} dz = -\frac{\zeta'(z)}{\zeta(z)} + \int \zeta'(z) d\left(\frac{1}{\zeta(z)}\right) = -\frac{\zeta'(z)}{\zeta(z)} - \int \left(\frac{\zeta'(z)}{\zeta(z)}\right)^2 dz =$$

$$(5) \qquad = \int_0^\infty u^{-z} \log u \; d\tilde{\omega}(u) - \int dz \left(\int_0^\infty u^{-z} \log u \; d\tilde{\omega}(u)\right)^2 =$$

$$= \int_0^\infty u^{-z} \log u \; d\tilde{\omega}(u) - \int dz \int_0^\infty u^{-z} \log u \; d\tilde{\omega}(u) \int_0^\infty v^{-z} \log v \; d\tilde{\omega}(v).$$

---

1) A. SELBERG, An elementary proof of the prime number theorem, *Annals of Math.*, (2) **50** (1949), pp. 305—313.

2) P. ERDŐS, On a new method in the elementary theory of numbers which leads to an elementary proof of the prime number theorem, *Proceedings National Academy of Sciences U.S.A.*, **35** (1949), pp. 374—384.

3) N. WIENER, Tauberian theorems, *Annals of Math.*, (2) **33** (1933), pp. 1 – 100.

Now,

$$-\int dz \int_0^\infty u^{-z} \log u \, d\tilde{\omega}(u) \int_0^\infty v^{-z} \log v \, d\tilde{\omega}(v) =$$

$$= -\int dz \int_0^\infty \log u \, d\tilde{\omega}(u) \int_0^\infty w^{-z} \log \frac{w}{u} \, d_v \tilde{\omega}\left(\frac{w}{u}\right) =$$

(6) $$= -\int dz \int_0^\infty \log u \, d\tilde{\omega}(u) \int_0^\infty w^{-z} d_w \int_0^{\frac{w}{u}} \log v \, d\tilde{\omega}(v) =$$

$$= -\int dz \int_0^\infty w^{-z} d_w \int_{u=0}^{u=\infty} \log u \, d\tilde{\omega}(u) \int_0^{\frac{w}{u}} \log v \, d\tilde{\omega}(v) =$$

$$= \text{const.} + \int_0^\infty w^{-z} \frac{1}{\log w} d_w \int_{u=0}^{u=\infty} \log u \, d\tilde{\omega}(u) \int_0^{\frac{w}{u}} \log v \, d\tilde{\omega}(v).$$

Hence,

(7) $$-\int \frac{\zeta''(z)}{\zeta(z)} dz = \text{const.} + \int_0^\infty u^{-z} \log u \, d\tilde{\omega}(u) +$$

$$+ \int_0^\infty u^{-z} \frac{1}{\log u} d_u \int_{w=0}^{w=\infty} \log w \, d\tilde{\omega}(w) \int_0^{\frac{u}{v}} \log v \, d\tilde{\omega}(v).$$

Let us now consider $\zeta(z)$ on the line $z = 1 + iy$. It is well known that $\zeta(z)$ is analytic on this line, except at $z = 1$, when it is of the form $\varphi(z) + \frac{1}{z-1}$, where $\varphi(z)$ is analytic. It consequently has on the line only zeros of finite order. Now, if ε is real and positive,

(8) $$\left|\frac{1}{\zeta(1+iy+\varepsilon)}\right| \leq \left|\sum \frac{\mu(n)}{n^{1+iy+\varepsilon}}\right| \leq \sum \frac{1}{n^{1+\varepsilon}} \leq \zeta(1+\varepsilon) = O\left(\frac{1}{\varepsilon}\right);$$

so that the zeta function cannot have a pole of higher order than 1 on the 1-line. It follows that $-\frac{\zeta'(z)}{\zeta(z)}$ cannot have a singularity on the 1-line that is not a pole of order 1 with residue -1, except for the singularity at $z = 1$, which is a pole of order 1 with residue 1. Thus either $-\frac{\zeta'(z)}{\zeta(z)}$ has no other singularity on the 1-line than the pole at $z = 1$, in which case

(9) $$\frac{1}{u} \int_0^u \log u \, d\tilde{\omega}(u) \to 1,$$

or it has a singularity at $1+i\lambda$. In the first case, we may write (9):

$$\frac{\log u\,\varpi(u)}{u}-\frac{1}{u}\int_0^u\frac{\varpi(u)\,du}{u}\to 1; \tag{10}$$

and since it is well known that $\varpi(u)=o(u)$, it follows that

$$\varpi(u)\sim\frac{u}{\log u}, \tag{11}$$

which is equivalent to the prime number theorem.

On the other hand, let

$$\zeta(1+i\lambda)=0. \tag{12}$$

Then

$$-\frac{\zeta'(z)}{\zeta(z)}-\frac{1}{2}\,\frac{\zeta'(z+i\lambda)}{\zeta(z+i\lambda)}-\frac{1}{2}\,\frac{\zeta'(z-i\lambda)}{\zeta(z-i\lambda)}= \tag{13}$$

$$=\int_0^\infty u^{-z}(1+\cos\lambda\log u)\log u\;d\varpi(u)$$

will be a Dirichlet series with non-negative coefficients with no singularity at $z=1$. It is easy to show that it can then have no singularity on $\Re z=1$. Thus by a very weak form of the Ikehara theorem,

$$\frac{1}{u}\int_0^u(1+\cos\lambda\log u)\log u\;d\varpi(u)\to 0. \tag{14}$$

This is to say that

$$\varpi(u)=\varpi_1(u)+\varpi_2(u), \tag{15}$$

where

$$\lim_{u\to\infty}\frac{1}{u}\int_0^u\log u\;d\varpi_1(u)=0, \tag{16}$$

and where $\varpi_2(u)$ only increases over the intervals

$$1+\cos\lambda\log u<\varepsilon, \tag{17}$$

or

$$|\lambda\log u-n\pi|<\varepsilon. \tag{18}$$

On the other hand, $-\int\frac{\zeta''(z)}{\zeta(z)}dz$ is a function which may be shown to behave like

$$\frac{\dfrac{2}{(z-1)^2}}{\dfrac{1}{z-1}}=\frac{2}{z-1} \tag{19}$$

at $z=1$, and to be analytic elsewhere in $x\geqq 1$, except for possible logarithmic singularities where $\zeta(z)=0$. These do not interfere with Lebesgue integrability.

Hence the Ikehara theorem applies, and

$$\text{(20)} \qquad \frac{1}{u}\int_0^u \log u \, d\tilde{\omega}(u) + \frac{1}{u}\int_0^u \frac{1}{\log u} d_u \int_{w=0}^{w=\infty} \log w \, d\tilde{\omega}(w) \int_0^{\frac{u}{w}} \log v \, d\tilde{\omega}(v) \to 2.$$

Hence

$$\text{(21)} \qquad \int_{|\lambda \log u - n\pi| < \varepsilon_1} \left(\log u \, d\tilde{\omega}(u) + \frac{1}{\log u} d_u \int_{w=0}^{w=\infty} \log w \, d\tilde{\omega}(w) \int_0^{\frac{u}{w}} \log v \, d\tilde{\omega}(v)\right) \sim$$
$$= 2e^{\frac{\pi n}{\lambda}}\left(e^{\frac{\varepsilon_1}{\lambda}} - e^{-\frac{\varepsilon_1}{\lambda}}\right),$$

so that

$$\text{(22)} \qquad \int_{|\lambda \log u - n\pi| < \varepsilon_1} \log u \, d\tilde{\omega}(u) \leqq 2e^{\frac{n\pi}{\lambda}}\left(e^{\frac{\varepsilon_1}{\lambda}} - e^{-\frac{\varepsilon_1}{\lambda}}\right).$$

Combining this with (16), (17), and (18), we see that for large u's

$$\text{(23)} \qquad \frac{1}{u}\int_0^u \log u \, d\tilde{\omega}(u) \leqq \frac{1}{u}\left(e^{\frac{\varepsilon_1}{\lambda}} - e^{-\frac{\varepsilon_1}{\lambda}}\right) \frac{2ue^{\varepsilon_1 \pi}}{1 - e^{-\pi/\lambda}},$$

which is asymptotically not greater than

$$\text{(24)} \qquad \frac{2\left(e^{\frac{\varepsilon_1}{\pi}} - e^{-\frac{\varepsilon_1}{\pi}}\right)}{1 - e^{-\pi/\lambda}}.$$

Since ε_1 is arbitrary, we see that

$$\text{(25)} \qquad \lim_{u \to \infty} \frac{1}{u}\int_0^u \log u \, d\tilde{\omega}(u) = 0.$$

This is however inconsistent with the known elementary Chebychev theorem, to the effect that

$$\text{(26)} \qquad \underline{\lim}_{u \to \infty} \frac{1}{u}\int_0^u \log u \, d\tilde{\omega}(u) > 0,$$

so that we have succeeded in eliminating the hypothesis that there is a λ for which $\zeta(1 + i\lambda) = 0$, and have proved the prime number theorem.

(Received July 11, 1949.)

Die Grenzschichte in der Theorie der gewöhnlichen Differentialgleichungen.

Von R. v. MISES in Cambridge, Mass.

Die Erscheinung der sogenannten „Grenzschichte“, die in verschiedenen Gebieten der klassischen Physik (Hydromechanik, Elastizität) eine Rolle spielt, beruht in letzter Linie auf einem mathematischen Tatbestand, der sich in seiner einfachsten Form in der Theorie der gewöhnlichen Differentialgleichungen aufweisen läßt. Vorgelegt sei die Randwertaufgabe zweiter Ordnung

(a) $$A(x, y)\, y'' + B(x, y)\, y' + C(x, y) = 0$$

(b) $$y(x_1) = y_1, \quad y(x_2) = y_2.$$

Man interessiert sich für das Verhalten der Lösung in dem Falle, daß der Faktor $A(x, y)$ im ganzen Bereich gleichmäßig gegen Null geht. Mit $A = 0$ wird (a) eine Gleichung erster Ordnung und hat im allgemeinen keine Lösung, die beide Bedingungen (b) befriedigt, wohl aber zwei verschiedene Lösungen $u(x)$ und $v(x)$ von denen jede eine der beiden Bedingungen erfüllt. Es stellt sich heraus, daß unter gewissen Voraussetzungen die Lösung $y(x)$ von (a), (b) im ganzen Bereich *mit Ausnahme eines schmalen Grenzstreifens* gegen $u(x)$ oder $v(x)$ konvergiert (je nach dem Vorzeichen von A) während der Übergang zur Erfüllung der anderen Randbedingung quasi sprungartig in der Grenzzone erfolgt.

Zur Illustration kann das Beispiel $B = C = 1$, $A = \nu = \text{const.}$ dienen. Hier läßt sich die Lösung von (a), (b) in zwei Formen anschreiben:

$$y(x) = y_2 + x_2 - x + (y_2 + x_2 - y_1 - x_1)\, \frac{e^{(x_1 - x)/\nu} - e^{(x_1 - x_2)/\nu}}{e^{(x_1 - x_2)/\nu} - 1} =$$

$$= y_1 + x_1 - x - (y_2 + x_2 - y_1 - x_1)\, \frac{e^{(x_2 - x)/\nu} - e^{(x_2 - x_1)/\nu}}{e^{(x_2 - x_1)/\nu} - 1}.$$

Wenn ν durch positive Werte gegen Null geht, wird der Bruch im ersten Ausdruck für $x - x_1 \geqq \varepsilon > 0$ gleich 0 und $y(x)$ gleich $y_2 + x_2 - x$. Nähert sich jedoch ν der Null von links, so hat der zweite Bruch den Grenzwert 0, sobald $x_2 - x \geqq \varepsilon > 0$, und dann $y(x) = y_1 + x_1 - x$.

Im folgenden wird das Auftreten der Grenzschicht unter der wesentlichen Annahme, daß A/B constantes Zeichen hat, bewiesen. Ein Zeichenwechsel von A/B verändert die Verhältnisse völlig.

1. Behauptungen.

Wir betrachten einen rechteckigen Bereich

(1) $$R:\ x_1 \leqq x \leqq x_2,\ b_1 \leqq y \leqq b_2,\ x_2 - x_1 = L,\ b_2 - b_1 = M.$$

Sei $P_1(x_1, y_1)$ ein Punkt der linken und $P_2(x_2, y_2)$ ein Punkt der rechten Begrenzung von R. Die Funktionen $f(x, y)$ und $g(x, y)$ seien stetig und beschränkt in R, überdies sei g positiv:

(2) $$|f(x, y)| \leqq F, \qquad 0 < g(x, y) \leqq G.$$

Schließlich werde angenommen, daß es zwei Funktionen $u(x)$ und $v(x)$ gibt, die durch

(3) $$u' = f(x, u),\ u(x_1) = y_1$$

(4) $$v' = f(x, v),\ v(x_2) = y_2$$

eindeutig im Intervall (x_1, x_2) definiert werden und ganz in R verlaufen. Unter diesen Voraussetzungen beweisen wir:

Satz 1. *Wenn die Randwertaufgabe*

(5) $$y' = f(x, y) + \nu g(x, y) y''$$

(6) $$y(x_1) = y_1,\ y(x_2) = y_2$$

bei kleinem positiven oder negativen ν eine Lösung $y(x)$ besitzt, so ist

(7) *bei* $\nu \geqq 0$: $\lim\limits_{\nu \to 0} y(x) = u(x)$ *gleichmäßig in* $x_1 \leqq x \leqq x_2 - \varepsilon \quad (\varepsilon > 0)$

(8) *bei* $\nu \leqq 0$: $\lim\limits_{\nu \to 0} y(x) = v(x)$ „ „ $x_1 + \varepsilon \leqq x \leqq x_2 \quad (\varepsilon > 0)$.

Der Beweis dieses Satzes stützt sich auf den folgenden Hilfssatz, der in gewisser Hinsicht mehr besagt:

Satz 2. *Sei $N = FL + M$. Wenn eine Integralkurve $y(x)$ von (5) die rechte Begrenzung von R trifft, d. h. wenn $b_1 \leqq y(x_2) \leqq b_2$, so gilt für genügend kleines $|\nu|$ im ganzen Intervall*

(9) $$|y'(x) - f(x, y)| \leqq \frac{2GN|\nu|}{(x_2 - x)^2}(1 + \varepsilon)$$

für jedes noch so kleine $\varepsilon > 0$. Ebenso, wenn $y(x)$ an der linken Begrenzung des Rechtecks beginnt, ist

(9′) $$|y'(x) - f(x, y)| \leqq \frac{2GN|\nu|}{(x - x_1)^2}(1 + \varepsilon).$$

Dieser Satz ist im wesentlichen gleichbedeutend mit dem folgenden:

Satz 3. *Sei $z(x) = y'(x) - f(x, y)$. Wenn an einer Stelle einer Integralkurve von (5) $z(x_0) > 0$, so gibt es zu jedem $\varepsilon > 0$ ein solches $\nu_1 > 0$, daß*

(10) $$z(x) \geqq (1 - \varepsilon) z(x_0) \quad \begin{cases} \textit{in } x_0 \leqq x \leqq x_2 & \textit{für } 0 < \nu < \nu_1, \\ \textit{in } x_1 \leqq x \leqq x_0 & \textit{für } 0 < -\nu < \nu_1. \end{cases}$$

Analog bei $z(x_0) < 0$.

2. Beweis der Sätze 3 und 2 für differenzierbares f.

Die beiden letzten Sätze — in etwas schärferer Form — lassen sich besonders leicht beweisen, wenn $f(x, y)$ als beschränkt differenzierbar vorausgesetzt wird. Wenn angenommen wird

$$|\operatorname{grad} f| \leqq K, \tag{11}$$

so gilt für Differentiation längs der Integralkurve

$$\left|\frac{df}{dx}\right| \leqq K(1+F). \tag{11'}$$

Wir wählen

$$\nu_1 = \frac{z(x_0)}{KG(1+F)}. \tag{12}$$

Dann folgt durch Differentiation von $z = y' - f$, solange $z > 0$, $\nu > 0$,

$$z' = y'' - \frac{df}{dx} \geqq \frac{z}{\nu_1 G} - \frac{df}{dx} \geqq K(1+F)\frac{z - z(x_0)}{z(x_0)}. \tag{13}$$

Diese Differential-Ungleichung für z besagt, daß bei $z(x_0) > 0$ die Differenz $z - z(x_0)$ mit wachsendem x nicht abnehmen kann, daß also $z - z(x_0) \geqq 0$ für $x > x_0$ und $z - z(x_0) \leqq 0$ für $x < x_0$. Damit ist die erste der vier Aussagen von Satz 3 bewiesen; die anderen sind analog.

Die Differentialgleichung (5) gibt für $z \geqq z(x_0) \gtrless 0$ und $\nu > 0$

$$y'' = \frac{z}{\nu g} \geqq \frac{z(x_0)}{\nu G} \tag{14}$$

woraus durch zweimalige Integration folgt

$$y - y_0 - y'(x_0)(x - x_0) \geqq \frac{z(x_0)}{2\nu G}(x - x_0)^2. \tag{15}$$

Wenn y für $x = x_2$ in das Intervall (b_1, b_2) fallen soll, ist die linke Seite von (15) dem Betrage nach nicht größer als $M + FL = N$, also muß

$$z(x_0) \leqq \frac{2G\nu}{(x_2 - x_0)^2} \tag{16}$$

sein, und dies ist die (verschärfte) Aussage (9) des Satzes 2.

3. Allgemeiner Beweis der Sätze 3 und 2.

Es sei $z(x_0) > 0$, $\nu > 0$. Wenn $f(x, y)$ stetig ist, so gibt es einen Kreis C_0, mit dem Mittelpunkt in $P_0(x_0, y_0)$ so daß in jedem Punkt innerhalb von C_0

$$|f(x, y) - f(x_0, y_0)| \leqq \varepsilon z(x_0). \tag{17}$$

Solange eine in P_0 beginnende Integralkurve von (5) innerhalb C_0 verläuft, kann y' nicht abnehmen. Denn wäre x die erste Abszisse, an der das geschieht, so müßte hier $y'(x) \geqq y'(x_0) = z(x_0) + f(x_0, y_0)$ und daher

$$y''(x) = \frac{y'(x) - f(x, y)}{\nu g(x, y)} \geqq \frac{z(x_0) + f(x_0, y_0) - f(x, y)}{\nu g(x, y)} \geqq \frac{(1-\varepsilon)\, z(x_0)}{\nu g(x, y)}$$

d. h. $y''(x)$ hat einen positiven Wert. Es ist also y' im ganzen Bereich C_0 nicht abnehmend und überall

$$y'' \geqq \frac{(1-\varepsilon)\, z(x_0)}{\nu G} = \alpha. \tag{18}$$

Somit liegt die Integralkurve, rechts von x_0, solange sie in C_0 verläuft, oberhalb der Parabel

$$y(x) = y_0 + y'(x_0)(x-x_0) + \frac{\alpha}{2}(x-x_0)^2. \tag{19}$$

Hierbei ist $y'(x_0) \geqq -F + z(x_0, y_0)$. Der Teil der Parabel (19), der zwischen dem Punkt $P_0(x_0, y_0)$ und einem Punkt $P_*(x_*, y_*)$ liegt, für den

$$\frac{y_* - y_0}{x_* - x_0}$$

einen vorgegebenen Wert A annimmt, kann durch Vergrößerung von α, also durch Verkleinerung von ν, auf das Innere des Kreises C_0 zusammengezogen werden. Wählt man ein entsprechend kleines ν, so erreicht die Integralkurve die Ordinate y_* an einer Abszisse $x' < x_*$ und es ist

$$\frac{y_* - y_0}{x' - x_0} > A$$

und a fortiori (da y' nicht abnehmend ist)

$$y'(x') > A. \tag{20}$$

Es läßt sich nun zeigen: Wenn in einem Punkt einer Integralkurve von (5) einmal $y' \geqq F + a$ bei positivem a, so ist von da an y' nicht abnehmend und $y'' \geqq a/\nu G$. Wäre nämlich x die erste Abszisse nach x', an der y' abnimmt, so müßte $y'(x) \geqq y'(x') \geqq F + a$ sein, daher

$$y''(x) = \frac{y'(x) - f(x, y)}{\nu g(x, y)} \geqq \frac{a + F - f(x, y)}{\nu g(x, y)} \geqq \frac{a}{\nu G} \tag{21}$$

im Widerspruch mit $y'' < 0$. Es ist also für alle $x > x'$ die Ungleichung (21) erfüllt.

Wir wählen nun die willkürlich gelassene Größe A so, daß

$$A = F + (1-\varepsilon)\, z(x_0). \tag{22}$$

Dann folgt aus (21) in Verbindung mit (18), daß $y'' \geqq \alpha$ im ganzen Intervall $x \geqq x_0$. Also bleibt die Integralkurve dauernd oberhalb der Parabel (19):

$$y(x) - y_0 - y'(x_0)(x - x_0) \geqq \frac{\alpha}{2}(x - x_0)^2.$$

Genau wie bei (15) beweist man daraus, daß

$$\alpha = \frac{(1-\varepsilon) z(x_0)}{\nu G} \leqq \frac{2N}{(x_2 - x_0)^2} \tag{23}$$

sein muß, was mit Gl. (9) von Satz 2 übereinstimmt.

Hinsichtlich Satz 3 beachte man, daß y' innerhalb C_0 nicht abnimmt und $f(x, y) \leqq f(x_0, y_0) + \varepsilon z(x_0)$ gilt, so daß innerhalb C_0

(24) $$z(x) = y'(x) - f(x, y) \geqq y'(x_0) - f(x_0, y_0) + \varepsilon z(x_0) \geqq (1-\varepsilon)\, z(x_0).$$

Von der Stelle (x', y_*) an ist $y' > A$, also nach (22) ebenfalls

(24′) $$z(x) = y'(x) - f(x,y) \geqq (1-\varepsilon) z(x_0) + F - f(x,y) \geqq (1-\varepsilon)\, z(x_0),$$

was zu beweisen war.

4. Beweis des Satzes 1.

Der Übergang von Satz 2 zu Satz 1 wird geliefert durch das bekannte Theorem der Theorie der Differentialgleichungen: Ist $|\varphi(x, y)|$ beschränkt im Intervall (x_1, a_2) und besitzt

$$y' = f(x, y) + \nu \varphi(x, y), \quad y(x_1) = y,$$

genau eine Lösung, für jedes ν, so konvergiert diese mit abnehmendem ν gegen die Lösung für $\nu = 0$, gleichmäßig in (x_1, a_2). In der Ungleichung (9) von Satz 2 ist der Faktor von ν beschränkt im Intervall $x_1 \leqq x \leqq x_2 - \varepsilon$, für beliebig kleines $\varepsilon > 0$. Ebenso ist der Faktor von ν an der rechten Seite von (9′) beschränkt im Intervall $x_1 + \varepsilon \leqq x \leqq x_2$. Damit ist Satz 1 bewiesen.

5. Asymptotisches Verhalten in der Grenzschicht.

Die übliche Betrachtungsweise der Grenzschicht-Theorie richtet sich nur auf das asymptotische Verhalten der Integrale in der Grenzschicht und sieht das im Vorstehenden bewiesene als gegeben an. Ist tatsächlich einmal der Verlauf der Lösung im Intervall $(x_1, x_2 - \varepsilon)$ bekannt, so läßt sich leicht der Rest bestimmen.

Wir beschränken uns auf den Fall $\nu > 0$, mit der Grenzschicht zwischen $x_2 - \varepsilon$ und x_2, und führen die zusätzliche Bezeichnung

(25) $$g(x, y) \geqq \gamma > 0$$

ein. Der Wert der in (3) definierten Funktion $u(x)$ an der Stelle x_2 heiße u_2.

Wird die Variable x ersetzt durch

(26) $$\xi = \frac{x_2 - x}{\nu},$$

so erhält die Differentialgleichung (5) die Form

(27) $$-\frac{dy}{d\xi} = g(x_2 - \nu\xi, y) \frac{d^2 y}{d\xi^2} + \nu f(x_2 - \nu\xi, y).$$

Da $|f|$ beschränkt und $1/g$ stetig ist, können wir den oben erwähnten Stetigkeitssatz (in seiner Verallgemeinerung auf Gleichungen zweiter Ordnung), anwenden. Mit $\nu = 0$ wird (27)

(28) $$\frac{d^2 y}{d\xi^2} = -\frac{1}{g(x_2, y)} \frac{dy}{d\xi}.$$

Mit der Abkürzung

$$h(y) = \int^{y} \frac{d\eta}{g(x_2, \eta)} \tag{29}$$

wird das Integral von (28), das für $\xi = 0$ den Wert y_2 und für $\xi = \infty$ den Wert u_2 annimmt, gegeben durch

$$\xi = \int_{y_2}^{y} \frac{d\eta}{h(u_2) - h(\eta)}.$$

Da die Ableitung von h zwischen $1/G$ und $1/\gamma$ liegt, geht in der Tat $h(u_2) - h(\eta)$ mit der ersten Potenz von $u_2 - \eta$ gegen Null. Mithin gibt

$$\int_{y_2}^{y} \frac{d\eta}{h(u_2) - h(\eta)} = \lim_{\nu \to 0} \frac{x_2 - x}{\nu} \tag{30}$$

den Verlauf von $y(x)$ in dem Grenzstreifen von $x_2 - \varepsilon$ bis x_2.

In der gleichen Weise läßt sich der Fall $\nu < 0$ erledigen.

HARVARD UNIVERSITY.

(Eingegangen am 17. August 1949.)

Fourier series with a sequence of positive coefficients.

By R. P. BOAS, JR. in Providence, R. I.

Let $f(\theta)$ be an integrable function of period 2π, with the Fourier series $\sum_{-\infty}^{\infty} c_n e^{in\theta}$. There are a number of theorems which indicate that if all c_n are nonnegative, then the magnitude of the c_n, and hence the behavior of $f(\theta)$, are controlled by the behavior of $f(\theta)$ near $\theta=0$. We mention two such results: if $c_n \geqq 0$ and $f(\theta)$ is bounded in a neighborhood of 0, then $\sum |c_n| < \infty$ and so $f(\theta)$ is continuous everywhere[1]; if $c_n \geqq 0$ and the pth derivative $f^{(p)}(\theta)$ exists at $\theta=0$, then $f^{(p-1)}(\theta)$ exists everywhere[2]. We shall generalize the latter theorem by assuming that the c_n are real and that the changes of sign in the sequence $\{c_n\}$ are not too frequent; our conclusion will be that $f(\theta)$ has at least a $(p-1)$th derivative everywhere if it has an integrable pth derivative in a sufficiently large neighborhood of $\theta=0$. Our theorem is, more precisely, as follows.

Theorem 1. *Let $f(\theta)$ be an integrable function of period 2π, with real Fourier coefficients c_n. Let k_n be the subsequence of integers at which a change of sign in the sequence c_n does not occur, and suppose that $|k_n - nB| < L$, where $B \geqq 1$ and L is a fixed positive number. Let p be a positive integer. If $f^{(p)}(\theta)$ exists and is integrable in the interval $(-\delta, \delta)$, where $\delta > \pi(1-B^{-1})$, then $f^{(p-1)}(\theta)$ exists for all θ. More generally, the same conclusion follows if, for some fixed positive A, we count c_n as "positive" if $c_n > -A|n|^{-p}$ and as "negative" if $c_n < A|n|^{-p}$, i. e., if we count no „change of sign" between c_n and c_{n+1} if either $c_n > -A|n|^{-p}$ and $c_{n+1} > -A|n+1|^{-p}$, or else $c_n < A|n|^{-p}$ and $c_{n+1} < A|n+1|^{-p}$.*

[1]) R. E. A. C. PALEY: see, for example, G. H. HARDY and W. ROGOSINSKI, *Fourier Series* (Cambridge, 1944), p. 72.

[2]) R. FORTET, Calcul des moments d'une fonction de répartition à partir de sa caractéristique, *Bulletin des Sciences Math.*, (2) 68 (1944), pp. 117—131; H. CRAMÉR, *Mathematical Methods of Statistics* (Princeton, 1946), p. 90.

Proof. We have

$$2\pi c_n=\int_{-\delta}^{\delta} f(t)e^{-int}dt+\int_{\delta}^{\pi} f(t)e^{-int}dt+\int_{-\pi}^{-\delta} f(t)e^{-int}dt,$$

$$(-1)^n 2\pi c_n=(-1)^n\int_{-\delta}^{\delta} f(t)e^{-int}dt+\int_{\delta}^{\pi} f(t)e^{-in(t-\pi)}dt+\int_{-\pi}^{-\delta} f(t)e^{-in(t+\pi)}dt=$$

$$=(-1)^n\int_{-\delta}^{\delta} f(t)e^{-int}dt+\int_{\delta-\pi}^{0} f(t+\pi)e^{-int}dt+\int_{0}^{\pi-\delta} f(t-\pi)e^{-int}dt=$$

$$=(-1)^n\left\{\frac{f(\delta)e^{-in\delta}-f(-\delta)e^{in\delta}}{-in}+\ldots+(-1)^{p-1}\frac{f^{(p-1)}(\delta)e^{-in\delta}-f^{(p-1)}(-\delta)e^{in\delta}}{(-in)^p}+\right.$$

$$\left.+\frac{(-1)^p}{(-in)^p}\int_{-\delta}^{\delta} f^{(p)}(t)e^{-int}dt\right\}+\int_{-(\pi-\delta)}^{\pi-\delta} g(t)e^{-int}dt,$$

where $g(t)=f(t-\pi\operatorname{sgn} t)$. Hence

$$n^p(-1)^n 2\pi c_n=\varphi(n)+\psi(n),$$

where

$$\varphi(n)=(-1)^n i^{-p}\int_{-\delta}^{\delta} f^{(p)}(t)e^{-int}dt,$$

and $\psi(n)$ is obtained by putting $z=n$ in

$$\psi(z)=\{iz^{p-1}[f(\delta)e^{iz(\pi-\delta)}-f(-\delta)e^{-iz(\pi-\delta)}]+\ldots+$$

$$+(-1)^{p-1}i^p[f^{(p-1)}(\delta)e^{iz(\pi-\delta)}-f^{(p-1)}(-\delta)e^{-iz(\pi-\delta)}]+z^p\int_{-(\pi-\delta)}^{\pi-\delta} g(t)e^{-izt}dt\}.$$

Thus $\varphi(n)=O(1)$ as $n\to\infty$ and $\psi(z)$ is an entire function such that $|\psi(z)|\leqq\text{const.}\,e^{(\pi-\delta)|z|}$. Furthermore, $\psi(x)$ is real for real x, as follows from our assumption that all c_n are real.

Consider now an index k_n. If both $c_{k_n}>-A|k_n|^{-p}$ and $c_{k_n+1}>-A|k_n+1|^{-p}$, and (for example) if k_n is even and positive, we have $\psi(k_n)\geqq-2\pi A-\varphi(k_n)\geqq-C$, and $\psi(k_n+1)\leqq 2\pi A-\varphi(k_n+1)\leqq C$, where C is some constant; hence $|\psi(m_n)|\leqq C$ for some number m_n, where $k_n\leqq m_n\leqq k_n+1$. If $k_{n+1}=k_n+1$, then $c_{k_{n+1}}$ satisfies the same inequality as c_{k_n} and c_{k_n+1}, and we determine m_{n+1} similarly, $k_n+1\leqq m_{n+1}\leqq k_n+2$. If possible we select m_n and m_{n+1} so that $m_{n+1}-m_n\geqq\frac{1}{2}$. If this is not possible, we must have $\psi(x)>C$ for $k_n\leqq x\leqq k_n+\frac{1}{2}$ and for $k_n+\frac{3}{2}\leqq x\leqq k_n+2$, for otherwise we could choose either $k_n\leqq m_n\leqq k_n+\frac{1}{2}$ and $k_n+1\leqq m_{n+1}\leqq k_n+2$, or $k_n\leqq m_n\leqq k_n+1$ and $k_n+\frac{3}{2}\leqq m_{n+1}\leqq k_n+2$. Then since $\psi(k_n+1)\leqq C$, it follows that $\psi(x)$ has a minimum between $k_n+\frac{1}{2}$ and $k_n+\frac{3}{2}$, and so m_n and m_{n+1} can be chosen so that they are separated by a point q_n such that $\psi'(q_n)=0$. On the other hand, if $k_{n+1}>k_n+1$, we have $k_n\leqq m_n\leqq k_n+1$, $m_{n+1}\geqq k_n+2$, and so certainly $m_{n+1}-m_n>\frac{1}{2}$. Similar considerations apply for odd or negative n, or when the inequality satisfied by c_{k_n} is reversed.

Thus $\psi(z)$ satisfies $|\psi(z)|\leqq \text{const.}\, e^{(\pi-\delta)|z|}$, and $|\psi(m_n)|\leqq C$, where $|m_n-nB|\leqq L+1$ and either $|m_{n+1}-m_n|\geqq\frac{1}{2}$ or else $\psi'(q_n)=0$ with q_n between m_n and m_{n+1}. Now if $\delta>\pi(1-B^{-1})$, we have $\pi-\delta<\pi/B$ and by a result of DUFFIN and SCHAEFFER[3]), $\psi(x)$ is bounded on the whole real axis. (DUFFIN and SCHAEFFER require $|m_{n+1}-m_n|>\gamma>0$, but an analysis of their proof shows that the theorem remains valid without this restriction if $\psi'(x)$ vanishes between any two m_n's which differ by less than some fixed γ.) Since $\psi(x)$ is bounded, in particular $\psi(n)$ is bounded, and since $\varphi(n)$ is bounded, $n^p c_n$ is bounded. Hence $\sum|n^{p-1}c_n|^2$ converges and $f(\theta)$ has a $(p-1)$th derivative (belonging to L^2).

There is an analogous theorem for power series which can be proved in a similar way (it would be possible to formulate a general result including both theorems):

Theorem 2. *Let* $F(z)=\sum_{n=0}^{\infty}a_n z^n$ *for* $|z|<1$ *and suppose that for* $-\delta\leqq\theta\leqq\delta$ *and* $0<r<1$, *we have* $|F(re^{i\theta})|\leqq\omega(\theta)$, *where* $\omega(\theta)$ *is integrable; let* $F(z)$ *have a radial boundary function* $F(e^{i\theta})$ *for* $-\delta\leqq\theta\leqq\delta$, *such that* $F(e^{i\theta})$ *has an integrable* pth *derivative in* $-\delta\leqq\theta\leqq\delta$. *Let the* a_n *be real and let* k_n *be the subsequence of positive integers at which a change of sign in the sequence* $\{a_n\}$ *does not occur. If* $|k_n-nB|\leqq L$, *where* $B\geqq 1$ *and* L *is fixed, then* $a_n=O(n^{-p})$ *and consequently* $F(z)$ *has a radial boundary function, with at least* $p-1$ *derivatives, for all* θ.

(Received July 17, 1949.)

[3]) R. J. DUFFIN and A. C. SCHAEFFER, Power series with bounded coefficients, *American Journal of Math.*, **67** (1945), pp. 141—154.

Deux exemples singuliers d'équations différentielles.

Par J. DIEUDONNÉ à Nancy.

Il est assez généralement connu que la plupart des théorèmes d'existence classiques de CAUCHY pour une équation différentielle

$$\mathbf{x}' = \mathbf{f}(t, \mathbf{x}) \tag{1}$$

où t est une variable réelle, $\mathbf{x}$ un vecteur dans un espace vectoriel E sur le corps $\mathbf{R}$, $\mathbf{f}$ une fonction continue dans un voisinage d'un point de $\mathbf{R}\times E$, à valeurs dans E, sont valables non seulement lorsque E est un espace de dimension *finie*, mais même lorsque E est un *espace de Banach* quelconque. Il y a toutefois deux propriétés classiques des équations (1) qui cessent d'être valables lorsque E n'est plus supposé de dimension finie, comme nous allons le montrer par des exemples.

1. La première de ces propriétés est le *théorème de Peano*, qui, lorsque $\mathbf{f}$ est *continue* au voisinage d'un point $(t_0, \mathbf{x}_0)$, affirme l'existence d'*au moins une intégrale* de (1) dans un voisinage de t_0, prenant la valeur $\mathbf{x}_0$ en ce point[1]).

Prenons pour E l'espace (c_0) de Banach, c'est-à-dire l'espace des suites $\mathbf{x} = (x_n)_{n\geqq 0}$ de nombres réels, telles que $\lim\limits_{n\to\infty} x_n = 0$, muni de la norme $\|\mathbf{x}\| = \sup\limits_n |x_n|$, qui en fait un espace de Banach. Pour tout $\mathbf{x} = (x_n)$ dans E, désignons par $\mathbf{y}$ la suite (y_n) définie par

$$y_n = \sqrt{|x_n|} + \frac{1}{n+1}.$$

Il est clair que $\lim\limits_{n\to\infty} y_n = 0$, donc $\mathbf{y} \in E$; si on pose $\mathbf{y} = \mathbf{f}(\mathbf{x})$, il résulte d'autre part de la continuité uniforme de la fonction $\sqrt{|x|}$ dans $\mathbf{R}$ que l'application $\mathbf{x} \to \mathbf{f}(\mathbf{x})$ est une application *continue* de l'espace E dans lui-même. Cependant, nous allons montrer que l'équation différentielle

$$\mathbf{x}' = \mathbf{f}(\mathbf{x}) \tag{2}$$

n'admet *aucune* intégrale *à valeurs dans E*, égale à 0 au point $t = 0$. En effet, si $\mathbf{u}(t)$ était une telle intégrale, on pourrait écrire $\mathbf{u}(t) = [u_n(t)]$, où u_n est une

---

[1]) Cf. E. KAMKE, *Differentialgleichungen reeller Funktionen* (Leipzig, 1930), p. 59.

fonction dérivable dans un voisinage de 0, satisfaisant à l'équation différentielle

$$u_n'(t) = \sqrt{|u_n(t)|} + \frac{1}{n+1} \tag{3}$$

et telle que $u_n(0) = 0$. Or, il est immédiat que chacune des équations (3) n'a qu'une seule intégrale satisfaisant à cette condition, comme on s'en assure par intégration effective; on vérifie ainsi que cette intégrale u_n est impaire et définie pour tout $t \in \mathbf{R}$, et que pour $t \geqq 0$, on a $u_n(t) \geqq \frac{t^2}{4}$. [2]) Cette dernière inégalité, valable pour tout n, montre aussitôt l'absurdité de notre hypothèse, car si $\mathbf{u}(t) \in E$ pour $t > 0$, il faut que la suite $[u_n(t)]$ tende vers 0 avec $1/n$ pour tout $t > 0$, ce qui n'est pas possible, comme nous venons de le voir.

2. Considérons maintenant une équation du type (1), où nous supposerons $\mathbf{f}$ définie et continue dans $I \times E$, où I est un intervalle d'extrémité finie t_0; nous supposerons en outre que dans $I \times E$, $\mathbf{f}$ est *localement lipschitzienne*, c'est-à-dire que pour tout point $(t, \mathbf{x})$ de $I \times E$, il existe un voisinage V de t, un voisinage W de $\mathbf{x}$ et un nombre $k > 0$ tels que $\|\mathbf{f}(s, \mathbf{x}_1) - \mathbf{f}(s, \mathbf{x}_2)\| \leqq k \|\mathbf{x}_1 - \mathbf{x}_2\|$ pour tout $s \in V$ et tout couple de points $\mathbf{x}_1, \mathbf{x}_2$ de W. Dans ces conditions, on démontre sans peine que pour tout $\mathbf{x}_0 \in E$, il existe un *plus grand* intervalle $J \subset I$, d'extrémité t_0, dans lequel existe une intégrale $\mathbf{u}$ de l'équation (1), à valeurs dans E et égale à $\mathbf{x}_0$ au point t_0, et cette intégrale est unique[3]). En outre, si $J \neq I$ et si la dimension de E est *finie*, on peut montrer que $\|\mathbf{u}(t)\|$ a une limite à droite égale à $+\infty$ à l'origine α de J. C'est cette dernière propriété qui, ainsi que nous allons le voir, ne subsiste plus lorsque E est de dimension *infinie*.

Nous prendrons pour E le même espace de Banach que dans le n°1. Désignons dans E par $\mathbf{e}_n$ la suite dont tous les termes sont égaux à 0 sauf le terme d'indice n, égal à 1; on peut alors écrire tout élément $\mathbf{x} = (x_n)$ de E sous la forme $\mathbf{x} = \sum_{n=0}^{\infty} x_n \mathbf{e}_n$. Posons, pour tout $\mathbf{x} = (x_n) \in E$

$$\mathbf{f}_n(\mathbf{x}) = [2(x_n + x_{n+1}) - 1]^+ (\mathbf{e}_{n+1} - \mathbf{e}_n).$$

Il est clair que $\mathbf{f}_n$ est continue et lipschitzienne dans E, que $\mathbf{f}_n(\mathbf{x}) = 0$ pour $\|\mathbf{x}\| \leqq \frac{1}{4}$ et $\mathbf{f}_n(\mathbf{x}) = \mathbf{e}_{n+1} - \mathbf{e}_n$ pour $\mathbf{x} = \lambda \mathbf{e}_n + (1-\lambda)\mathbf{e}_{n+1}$ (λ scalaire quelconque). D'autre part, pour tout entier $n > 0$, soit φ_n une fonction numérique $\geqq 0$, définie et continue dans l'intervalle $\frac{1}{n+1} \leqq t \leqq \frac{1}{n}$, égale à 0 aux extrémités de cet intervalle, et telle que $\int_{\frac{1}{n+1}}^{\frac{1}{n}} \varphi_n(t)\,dt = 1$. Soit I l'intervalle ouvert $t \leqq 1$ dans $\mathbf{R}$; définissons dans $I \times E$ la fonction suivante $\mathbf{f}$, à valeurs dans E:

pour $t \leqq 0$, $\mathbf{x}$ quelconque dans E, $\mathbf{f}(t, \mathbf{x}) = 0$;

---

[2]) Cela résulte aussi de théorèmes généraux sur les équations différentielles (cf. Kamke, *loc. cit.*, p. 81—85).

[3]) Pour le cas où E est de dimension finie, voir Kamke, *loc. cit.*, p. 135—136.

$$\text{pour } \frac{1}{n+1} \leqq t \leqq \frac{1}{n} \text{ et } \mathbf{x} \in E, \quad \mathbf{f}(t, \mathbf{x}) = \varphi_n(t)\mathbf{f}_n(\mathbf{x}) \qquad (n \geqq 1).$$

Il est immédiat que $\mathbf{f}$ est continue et localement lipschitzienne en tout point $(t_0, \mathbf{x}_0)$ de $I \times E$ tel que $t_0 \neq 0$; montrons qu'il en est encore de même en un point de la forme $(0, \mathbf{a})$, où $\mathbf{a} = (a_n)$ est quelconque dans E. En effet, il existe par hypothèse un entier m tel que, pour $n \geqq m$, on ait $|a_n| \leqq \frac{1}{8}$; pour tout $\mathbf{x} = (x_n)$ tel que $\|\mathbf{x} - \mathbf{a}\| \leqq \frac{1}{8}$, on a donc $|x_n| \leqq \frac{1}{4}$ pour $n \geqq m$, et par suite $\mathbf{f}_n(\mathbf{x}) = 0$ pour tout $n \geqq m$; on a donc $\mathbf{f}(t, \mathbf{x}) = 0$ pour $0 \leqq t \leqq 1/m$ et $\|\mathbf{x} - \mathbf{a}\| \leqq \frac{1}{8}$, ce qui établit notre assertion.

Définissons maintenant une suite de fonctions $\mathbf{v}_n$ à valeurs dan E, des la façon suivante:

$$\mathbf{v}_1(t) = 0 \text{ pour } t < \tfrac{1}{2}; \quad \mathbf{v}_1(t) = \mathbf{e}_1 + (\mathbf{e}_2 - \mathbf{e}_1) \int_t^1 \varphi_1(s)\,ds \text{ pour } \tfrac{1}{2} \leqq t \leqq 1;$$

$$\mathbf{v}_n(t) = 0 \text{ pour } t > \frac{1}{n},\ t < \frac{1}{n+1}; \quad \mathbf{v}_n(t) = (\mathbf{e}_{n+1} - \mathbf{e}_n) \int_t^{1/n} \varphi_n(s)\,ds \text{ pour } \frac{1}{n+1} \leqq t \leqq \frac{1}{n}.$$

Pour tout t tel que $0 < t \leqq 1$, posons $\mathbf{u}(t) = \sum_{n=1}^{\infty} \mathbf{v}_n(t)$, somme qui a toujours un sens puisqu'elle ne comporte qu'un nombre fini de termes non nuls; pour $\frac{1}{n+1} \leqq t \leqq \frac{1}{n}$, on a $\mathbf{u}(t) = \mathbf{e}_n + (\mathbf{e}_{n+1} - \mathbf{e}_n) \int_t^{1/n} \varphi_n(s)\,ds$ d'où résulte aussitôt que $\|\mathbf{u}(t)\| \leqq 1$ pour $0 < t \leqq 1$, que $\mathbf{u}$ est dérivable en tout point de cet intervalle et qu'on a $\mathbf{u}'(t) = -\mathbf{f}[t, \mathbf{u}(t)]$ en vertu de la définition de $\mathbf{f}$; enfin, comme $\mathbf{u}(1/n) = \mathbf{e}_n$, $\mathbf{u}(t)$ ne tend vers aucune limite lorsque t tend vers 0, ce qui montre que 0 est l'origine du plus grand intervalle J d'extrémité 1 dans lequel existe une intégrale de $\mathbf{x}' = -\mathbf{f}(t, \mathbf{x})$ qui prend la valeur $\mathbf{e}_1$ au point $t = 1$; on a donc $J \neq I$, et $\mathbf{u}(t)$ reste *bornée* dans J, ce que nous voulions établir.

(Reçu le 22 août 1949)

Bilinear functionals over $C \times C$.

By MARSTON MORSE in Princeton, N. J.

§ 1. F. Riesz's theorem and the Fréchet generalization.

It is perhaps appropriate in this volume in honor of F. RIESZ and L. FEJÉR that a summary account be given of new and unpublished theorems on the representation and uses of bilinear functionals over the cartesian product $C \times C$. These results were recently obtained by the author and Dr. WILLIAM TRANSUE. The famous theorem of F. RIESZ on the representation of the most general functional f, linear over the Banach space C, as a Riemann-Stieltjes integral

$$f(x) = \int_0^1 x(s)\,dg(s) \qquad [x \in C] \tag{1.1}$$

(where g is a functional of bounded Jordan variation over the interval [0,1]) was followed by FRÉCHET's representation of the most general functional Φ bilinear (including continuous) over the cartesian product $C \times C$. FRÉCHET represents Φ by a repeated Riemann—Stieltjes integral

$$\Phi(x, y) = \int_0^1 x(s)\,d_s \int_0^1 y(t)\,d_t k(s, t) \qquad [x, y \in C]. \tag{1.2}$$

The distribution function k was required to have a special finite variation $P(E, k)$ (here termed an F-variation) over the unit interval $E = [0, 1] \times [0, 1]$ on which k was defined. Bearing in mind the celebrated contributions of FEJÉR to the theory of Fourier series I am happy to include in this account the innovation in the theory of the Pringsheim convergence of double Fourier series which our new theorems on the nature of the F-variation make possible. This report will be restricted to bilinear as distinguished from multilinear functionals, and to the F-variation over the 2-dimensional intervals I as distinguished from the F-variation over the corresponding n-dimensional interval $I^{(n)}$. The major part of our theorems have, however, been extended to the n-dimensional case (see MT 6). There remain outstanding difficulties which have been solved only for the case $n = 2$.

We shall recall the definition of $P[I, k]$ and extend this definition. Let E' and E'' respectively represent the interval $[0, 1]$ on the s and t-axes. We admit a partition π of $E = E' \times E''$ into subintervals defined by straight lines $[s = s_r]$ and $[t = t_n]$. The values s_r and t_n used to define π shall satisfy the conditions

(1.3) $$0 = s_0 < s_1 < \dots < s_{r(\pi)} = 1,$$

(1.4) $$0 = t_0 < t_1 < \dots < t_{n(\pi)} = 1.$$

For $r = 1, \dots, r(\pi)$; $n = 1, \dots, n(\pi)$ set

(1.5) $$\Delta_{rn}(k) = k(s_r, t_n) - k(s_{r-1}, t_n) - k(s_r, t_{n-1}) + k(s_{r-1}, t_{n-1}).$$

Let e'_r be a constant, with $|e'_r| \leq 1$, associated with the rth interval of the partition (1.3) of E', and let e''_n with $|e''_n| \leq 1$ be similarly associated with the nth interval of the partition (1.4) of E''. We say that the set

(1.6) $$[e'_1, \dots, e'_{r(\pi)};\ e''_1, \dots, e''_{n(\pi)}] = e$$

is *associated* with the partition π.

Then by definition

(1.7) $$P(E, k) = \sup_{\pi, e} \sum_{r, n} e'_r e''_n \Delta_{rn}(k)$$

taking the sup over all admissible partitions π of E and associated sets e. We admit the possibility that $P(I, k) = +\infty$.

One immediately extends the definition of $P[I, k]$ over any closed subinterval $I = U \times V$ of E, restricting the partitioning values s_r to the interval U of the s-axis and the partitioning values t_r to the interval V of the t-axis. We also extend the definition of $P[I, k]$ to the case in which U and V may be open at either end point, both end points, or neither end point. In this case we set

(1.8) $$P[I, k] = \sup_{J} P[J, k]$$

where J ranges over all closed subintervals of I. (MT 6, § 2.)

So defined $P[I, k]$ should be compared with the Vitali-variation $V[I, k]$. This may be defined as

(1.9) $$V[I, k] = \sup_{\pi} \sum_{r, n} |\Delta_{rn}(k)|.$$

in case I is closed, and as $\sup\limits_{J} V(J, k)$, taken as in (1.8), in case I is not closed. It is immediately obvious that

(1.10) $$P[I, k] \leq V[I, k].$$

It was known previously that functions k exist for which $P(I, k) < \infty$ and $V[I, k] = \infty$, but the example given was of a function k which vanished almost everywhere in E, and was inadequate for the purposes of our theory. (See Clarkson and Adams.) How much less restrictive numerical conditions on k

in terms of the F-variation are, than corresponding conditions in terms of the V-variation, is shown by the following theorem, established in MT 8, § 6.

Theorem 1.1. *Let I_{xy} be the interval $[0, x] \times [0, y]$, and let p be an arbitrary positive number. There exists a function k mapping E continuously into the axis of reals, vanishing on the boundary of E, absolutely continuous in the sense of Carathéodory over every closed subinterval of $(0, 1) \times (0, 1)$ and such that*

$$V[I_{xy}, k] = \infty, \quad P[I_{xy}, k] \leqq x^p y^p \tag{1.11}$$

for arbitrary positive x and y.

There are other important respects in which the F-variation differs from the V-variation. If I and J are 2-intervals intersecting in a common edge then

$$V[I \cup J, k] = V[I, k] + V[J, k] \tag{1.12}$$

while

$$P[I \cup J, k] \leqq P[I, k] + P[J, k] \tag{1.13}$$

with the equality in general not holding in (1.13). In addition the decomposition $k = P - N$ of k into two monotone functions P and N, possible when $V(I, k) < \infty$, is not in general possible when $P(I, k) < \infty$. In spite of these considerable differences the F-variation can be used with great advantage in place of the V-variation in many branches of analysis.

§ 2. Some basic properties of the F-variation.

The properties of the F-variation described in this section parallel in a remarkable way well known properties of the Jordan variation $T_a^b(g)$ over the interval $[a, b]$ of a function g with values $g(s)$ defined for $s \in [a, b]$. Assuming that $T_0^1(g) < \infty$ we list the known properties to which we shall give analogies for the F-variation.

I. The limits $g(s-)$ and $g(s+)$ exist for $s \in (0, 1]$ and $[0, 1)$ respectively.

II. The points at which g fails to be continuous are at most countably infinite.

III. If g^+ and g^- are functions defined by setting $g^+(s) = g(s+)$, $g^-(s) = g(s-)$ for $s \in (0, 1)$ and $g(0) = g^+(0) = g^-(0)$, $g(1) = g^+(1) = g^-(1)$ then

$$T_0^1(g^+) = T_0^1(g^-) \leqq T_0^1(g).$$

IV. Referring to the Riesz representation (1.1) we have

$$\sup_{x \in C} \frac{|f(x)|}{\|x\|} = T_0^1(g^+) = T_0^1(g^-) \quad (\|x\| \neq 0).$$

V. If $c < s < s' < 1$ then for fixed c and variable s and s' $\lim\limits_{s' \to c} T_s^{s'}(g) = 0$.

The analogies of these properties of the Jordan variation have been obtained in MT 1, 2, 6 and 10. We shall suppose that $P[E, k] < \infty$ and that on at least one section K of E on which $s = \text{const.}$, and on one section of E on which $t = \text{const.}$, the function $k|K$ defined by k over K has a finite Jordan variation. We say then that k is in $\widehat{F}(E)$. For $k \in \widehat{F}(E)$ the following holds:

I. *Let (a, b) be an arbitrary point in the (s, t) plane and let S_{ab} be any one of the four open quadrants into which the (s, t)-plane is divided by the lines $s = a$ and $t = b$. For fixed (a, b) and for $(s, t) \in S_{ab}$*

$$\lim_{(s,t)\to(a,b)} k(s, t) = \bar{k}^S(a, b) \qquad [(s, t) \in S_{ab}] \tag{2.1}$$

exists whenever S_{ab} intersects E. The four limits corresponding to the four quadrants may all be different. (MT 1, Theorem 5.1.)

II. *The points in E at which k fails to be continuous lie on a countable number of straight lines parallel to the coordinate axes.* (MT 1, Theorem 6.3.)

III. *Corresponding to any one of the variable quadrants S_{ab} of I taken with a fixed orientation, we shall define a function k^S over E. The detailed definition in case $s > a$ and $t > b$ in S_{ab} follows. Let $k^S(a, b) = \bar{k}^S(a, b)$ for $(a, b) \in (0, 1) \times (0, 1)$, with $k^S(a, b) = k(a, b)$ at each corner point of E. Let $k^S(0, t) = k(0, t+)$ and $k^S(1, t) = k(1, t+)$ for $t \in (0, 1)$; and let $k^S(s, 0) = k(s+, 0)$, $k^S(s, 1) = k(s+, 1)$ for $s \in (0, 1)$. The remaining three functions k^S are similarly defined. Then $k^S(s, t) = k(s, t)$ at each point of continuity of k and for any two quadrant types S and S', of fixed orientation,*

$$P[E, k^S] = P[E, k^{S'}] \leqq P[E, k].$$

(MT 6, Theorem 8.2, § 6.)

IV. *Referring to the Fréchet representation* (1.2)

$$\sup_{x,y} \frac{|\varphi(x, y)|}{\|x\|\,\|y\|} = P[E, k^S] \qquad [0 \neq x \in C; 0 \neq y \in C]$$

where S is any one of the four quadrant types of fixed orientation. (MT 2, Theorem 12.1, MT 3, § 3.)

V. *Let S_{ab} be an open quadrant with vertex at $(a, b) \in E$. Let J be a variable 2-interval in $S_{ab} \cap E$. Then $P[J, k] \to 0$ as the maximum distance of the vertices of J from an edge of S_{ab} tends to zero.* (MT 10, Theorem 3.1.)

Property V has the following important application. If k is continuous over E and in $\widehat{F}(E)$, then $P[J, k] \to 0$ uniformly for arbitrary choice of $J \in E$ as the area of J tends to zero. (See MT 10, Corollary 3.5.)

We add the useful fact that when k is in $\widehat{F}(E)$, k is bounded and measurable over E. (See MT 4, § 2.)

§ 3. The Pringsheim convergence of double Fourier series.

On observing that the Dirichlet integral

$$(3.1)\qquad \frac{1}{\pi^2}\int_0^\pi\int_0^\pi f(s,t)\,\frac{\sin\left(m+\frac{1}{2}\right)s}{\sin\frac{s}{2}}\;\frac{\sin\left(n+\frac{1}{2}\right)t}{\sin\frac{t}{2}}\,ds\,dt$$

which gives the partial sum S_{mn} of the Fourier series for f, is really a special evaluation $\Phi(x, y)$ of a bilinear functional Φ with

$$x(s)=\frac{\sin\left(m+\frac{1}{2}\right)s}{\sin\frac{s}{2}},\quad y(t)=\frac{\sin\left(n+\frac{1}{2}\right)t}{\sin\frac{t}{2}},$$

it is clear that the general theory of such functionals should be relevant to the convergence problem. Just as the second law of the mean is historically associated with the 1-dimensional Dirichlet integral, so here there is a generalized law of the mean which involves the F-variation of f, and is fundamental in treating the Dirichlet integral. As established in MT 10, Theorem 5.1, this law may be stated as follows.

Theorem 3.1. *Let f_1 and f_2 be integrable over* $[0, 1]$. *Let I the open interval $I=(0,1)\times(0,1)$. Suppose $g\in\widehat{F}(I)$ and that $g(0+,t)=0$ for $t\in(0,1)$ and $g(s,0+)=0$ for $s\in(0,1)$. Then*

$$(3.2)\qquad \left|\iint_I f_1(s)f_2(t)g(s,t)\,ds\,dt\right| \leqq \left|\int_c^1 f_1(s)\,ds\right|\left|\int_d^1 f_2(t)\,dt\right| P[I,g]$$

for some choice of c and d in $[0, 1]$.

With the aid of this theorem one can extend the 1-dimensional Jordan test as follows.

Theorem 3.2. *Let f be integrable over the interval $J=(0,2\pi)\times(0,2\pi)$, have the period 2π in each of its arguments, and be in $\widehat{F}(J)$. Then the Fourier series for f converges at each point (a, b) to the mean of the four open quadrant limits of f at (a, b). If f is continuous this convergence is uniform over J.* (MT 7, Theorem 1.)

As Theorem 1.1 indicates, the class of functions f which sastisfy this test is much larger than the class defined by Hardy in extending the Jordan test, using the V-variation of f. This theorem of course implies the Hardy theorem but not conversely.

We have weakened all of the classical two-dimensional tests (known to us) which use V-variations. These tests appear to include all the tests for Pringsheim convergence except the Tonelli test, and we prove that the Tonelli test is more restrictive than some of our "MT-tests". The tests so modified include those generalizing the Jordan, Dini, Young-Pollard, de la Vallée Poussin, Lebesgue and Gergen tests. (See MT 11.) We shall refer to the Dini test as typical.

If one is concerned with convergence at the origin one sets

$$4F(s,t)=f(s,t)+f(-s,t)+f(s,-t)+f(-s,-t)$$

and then writes $\varphi(s,t)=F(s,t)-c$. In the Young-Dini test the Fourier series of f converges to c if

$$\int_0^\pi\int_0^\pi\left|\frac{\varphi(s,t)}{st}\right|ds\,dt<\infty. \tag{3.3}$$

A much weaker version of this condition may be obtained as follows.

Let J be the interval $(0,\pi)\times(0,\pi)$. If g is defined over J and integrable over every closed subinterval of J, set

$$\bar{g}(u,v)=\int_\pi^u\int_\pi^v g(s,t)\,ds\,dt \qquad [(u,v)\in J].$$

The condition (3.3) may be written in the form

$$V[J,\overline{(\varphi/st)}]<\infty \qquad [\varphi=F-c] \tag{3.4}$$

and our weakened form of the Dini test is as follows.

Theorem 3.3. *If φ is integrable over $J=(0,\pi)\times(0,\pi)$, and if*

$$P[J,\overline{(\varphi/st)}]<\infty, \tag{3.5}$$

then the Fourier series of f converges to c at the origin.

This MT-Dini test is less restrictive than the Young-Dini test, as we show by an example. It is in fact so broad in its coverage that the class of functions φ which satisfy its conditions are included in none of the classical tests, not excepting the Gergen test (MT 11, § 11). We add the following

Theorem 3.4. *Each MT-test is less restrictive than the corresponding classical test.* (See Gergen for enumeration of classical tests.) *In particular the MT-test modifying the Gergen test is less restrictive than each classical test for Pringsheim convergence.* (See MT 1.1, Theorem 1.1.)

The last result is somewhat surprising since there is no proof known to us that the Gergen test in its original form is actually less restrictive than some of the other tests of Lebesgue type in the form given by Gergen. Our MT-Gergen test is, however, shown to be less restrictive than each of these tests.

Among the many theorems necessary in the calculus of Fréchet variations we shall state three which are typical.

Theorem 3.5. *Suppose g is integrable over every closed subinterval of $J=(0,1)\times(0,1)$ and that $P[J,\bar{g}]<\infty$. Let f_1 and f_2 be two functions in the Banach space M of functions essentially bounded over* $[0,1]$. *Then the function f_1f_2g with values $f_1(s)f_2(t)g(s,t)$ over J satisfies the condition*

$$P[J,\overline{(f_1f_2g)}]\leqq\|f_1\|\,\|f_2\|\,P[J,\bar{g}] \tag{3.6}$$

where $\|f_1\|$ and $\|f_2\|$ are norms of f_1 and f_2 in M. (MT 10, Theorem 6.2).

We point out that g may not be integrable over J so that it is possible that $V(J, \bar{g}) = \infty$. In $\bar{g}$ we really have a Harnack integral with special properties. Under the conditions of the theorem one can show that $\bar{g}$ has a continuous extension over $\bar{J}$; we term this extension of $\bar{g}$ over $\bar{J}$ an FL-integral of g and develop its properties (MT 10, § 6). If $V[J, \bar{g}] = \infty$, then for g in the theorem

$$P[J, \bar{g}] < \infty, \qquad P[J, \overline{|g|}] = V[J, \bar{g}] = \infty.$$

Thus g may have an FL-integral $\bar{g}$ over $\bar{J}$ while $|g|$ is not FL-integrable.

We refer to a theorem of Lebesgue. Let $x > 0, y > 0$ be positive infinitesimals. If g is in L over $J = (0, 1) \times (0, 1)$, according to Lebesgue

$$\int_0^{1-x} \int_0^{1-y} |\varDelta_{xy} g(s, t)| ds dt = o(1), \tag{3.7}$$

where $\quad \varDelta_{xy} g(s, t) = g(s+x, t+y) - g(s, t+y) - g(s+x, t) + g(s, t)$.

If J^{xy} denotes the interval of integration in (3.7), one can write (3.7) in the form

$$V[J^{xy}, \overline{\varDelta_{xy} g}] = o(1). \tag{3.8}$$

Theorem 3.6. *If $\bar{g}$ is an FL-integral over $\bar{J} = [0, 1] \times [0, 1]$ then*

$$P[J^{xy}, \overline{\varDelta_{xy} g}] = o(1). \tag{3.9}$$

(MT 10, Theorem 8.1.)

As we have shown, (3.9) can hold without g being integrable over $\bar{J}$, or (3.7) holding. (MT 8, § 6.)

When g is integrable over $J = (0, 1) \times (0, 1)$, the mean $\widehat{g}/uv$, where

$$\widehat{g}(u, v) = \int_0^u ds \int_0^v g(s, t) dt \qquad [(u, v) \in J],$$

enters frequently, for example in the generalized de la Vallée Poussin test. The following theorem is then useful. See MT 10, Theorem 7.1 and Corollary 7.1.

Theorem 3.7. *If g is in L over $J = (0, 1) \times (0, 1)$ and $f_1(s)$ and $f_2(t)$ are positive, monotone decreasing and continuous over* $(0, 1)$, *then*

$$P[J, f_1 f_2 \widehat{g}] \leq 4\, P[J, \overline{f_1 f_2 g}], \tag{3.10}$$

$$P[J, (\widehat{g}/st)] \leq P[J, g]. \tag{3.11}$$

In this theorem the right member of (3.10) can be finite while $f_1 f_2 g$ is not integrable over J. (MT 8, § 6.)

§ 4. Variational theory.

The preceding results are a by-product of the studies of Morse and Transue in the variational theory of quadratic functionals. For the purposes of this variational theory one considers the product $A \times B$ of any two normed linear vector spaces A and B of which the products

$$L_2 \times L_2,\ L_2 \times L_1,\ C \times L_1,\ C \times L_2,\ C \times C$$

are important special cases. Let us suppose here that the spaces A and B are spaces of functions defined on the interval [0, 1]. Given any distribution function k mapping the interval $E = E' \times E''$ into the axis of reals, a variation $h(A, B, k)$ of k over E is defined generalizing the definition of $P[E, k]$. The Riesz—Fréchet representation theory is then extended and a variational theory and spectral theory initiated. This spectral theory is not a transformation theory but rather a direct critical point theory with the aspects characteristic of such a theory.

Bibliography.

Banach, S., *Théorie des opérations linéaires.* (Warsaw, 1932).

Clarkson, J. A. and Adams, C. R., On definitions of bounded variation for functions of two variables, *Transactions American Math. Society,* **35** (1933), pp. 824–854.

Fréchet, M., Sur les fonctionnelles bilinéaires, *Ibidem,* **16** (1915), pp. 215—234.

Gergen, J. J., Convergence criteria for double Fourier series, *Ibidem,* **35** (1933), pp. 29—63.

Hardy, G. H., On double Fourier series, and especially those which represent the double zeta-function with real and incommensurable parameters, *Quarterly Journal of Math.,* **37** (1906), pp. 53—79.

Morse, M. and Transue, W., 1. Functionals of bounded Fréchet variation, *Canadian Journal of Math.,* **1** (1949), pp. 153—165. — 2. Functionals F bilinear over the product $A \times B$ of two p-normed vector spaces. I. The representation of F, *Annals of Math.,* **50** (1949). — 3. Functionals F bilinear over the product $A \times B$ of two p-normed vector spaces. II. Admissible spaces A, *Annals of Math.,* **51** (1950). — 4. Integral representations of bilinear functionals, *Proceedings National Academy of Sciences U. S. A.,* **35** (1949), pp. 136—143. — 5. A characterization of the bilinear sums associated with the classical second variation, *Annali di Matematica,* **29** (1950). — 6. The Fréchet variation in the small, sector limits, and left decompositions (To be published). — 7. The Fréchet variation and a generalization for multiple Fourier series of the Jordan test. (To be published). — 8. Norms of distribution functions associated with bilinear functionals. (To be published). — 9. The Fréchet variation and the convergence of multiple Fourier series, *Proceedings National Academy of Sciences U. S. A.,* **35** (1949), pp. 395—399. — 10. Calculus for Fréchet variations. (To be published). — 11. The Fréchet variation and Pringsheim convergence of double Fourier series. (To be published).

Riesz, F., Sur les opérations fonctionnelles linéaires, *Comptes Rendus de l'Académie des Sciences Paris,* **149** (1909), pp. 974—977.

Institute for Advanced Study.

(Received July 29, 1949.)

Solution de l'équation $\omega^{\xi} = \xi^{\omega}$ pour les nombres ordinaux.

Par Wacław Sierpiński à Varsovie.

On sait qu'il n'existe qu'un seul système α, β de nombres ordinaux finis α et $\beta > \alpha$, tels que $\alpha^{\beta} = \beta^{\alpha}$ (à savoir $\alpha = 2$, $\beta = 4$). Or, il est intéressant qu'il existe une infinité de systèmes α, β de nombres ordinaux transfinis, tels que $\alpha < \beta$ et $\alpha^{\beta} = \beta^{\alpha}$, même tels où $\alpha = \omega$. Le but de cette note est de les trouver tous. En d'autres termes, nous nous proposons de trouver tous les nombres ordinaux $\xi > \omega$ satisfaisant à l'équation

$$(1) \qquad \omega^{\xi} = \xi^{\omega}.$$

Il est à remarquer que l'équation (1) n'a pas de solutions en nombres ordinaux $\xi < \omega$ En effet, le nombre $\xi = 1$ ne satisfait pas évidemment à l'équation (1). Donc, si $\xi < \omega$ et $\omega^{\xi} = \xi^{\omega}$, on a $\xi = n$, où n est un nombre naturel > 1, et alors on a, comme on sait $n^{\omega} = \omega$, d'où $\omega^{n} \geqq \omega^{2} > \omega = n^{\omega}$, donc, vu que $n = \xi$, $\omega^{\xi} > \xi^{\omega}$, ce qui implique une contradiction.

Soit $\xi > \omega$ une solution de l'équation (1). Soit $\xi = \omega^{\alpha} a + \omega^{\beta} b + \dots$ le développement normal du nombre ξ. Vu que $\xi > \omega$, on a ici $\alpha \geqq 1$ et $\xi < \omega^{\alpha+1}$, d'où $\xi^{\omega} \leqq \omega^{(\alpha+1)\omega} = \omega^{\alpha\omega}$ (puisque, quel que soit le nombre ordinal α, on a $(\alpha+1)\omega = \alpha\omega$). D'autre part, comme $\xi \geqq \omega^{\alpha}$ (vu que a est un nombre naturel), on a $\xi^{\omega} \geqq \omega^{\alpha\omega}$. On a ainsi $\xi^{\omega} = \omega^{\alpha\omega}$ et, d'après (1), on trouve $\omega^{\xi} = \omega^{\alpha\omega}$, ce qui donne $\xi = \alpha\omega$. Comme $\xi > \omega$, il en résulte que $\alpha \geqq \omega$ (puisque, pour α naturels, on a $\alpha\omega = \omega < \xi$).

Soit $\alpha = \omega^{\gamma} l + \dots$ le développement normal du nombre α; vu que $\alpha \geqq \omega$, on a ici $\gamma \geqq 1$ et $\xi = \alpha\omega = \omega^{\gamma+1}$. Or, comme $\xi = \omega^{\alpha} a + \dots$ (le développement normal d'un nombre ordinal étant unique), on trouve $\alpha = \gamma + 1$, d'où $\alpha\omega = (\gamma+1)\omega = \gamma\omega$ et, comme $\alpha\omega = \omega^{\gamma+1}$, on trouve $\gamma\omega = \omega^{\gamma+1}$. Si $\gamma = \omega^{\mu} m + \dots$ est le développement normal du nombre γ, on a $\gamma\omega = \omega^{\mu+1}$, donc, vu que $\gamma\omega = \omega^{\gamma+1}$, on trouve $\omega^{\gamma+1} = \omega^{\mu+1}$, d'où $\gamma + 1 = \mu + 1$ et $\gamma = \mu$. D'après $\gamma = \omega^{\mu} m + \dots$ on trouve donc $\gamma \geqq \omega^{\mu} = \omega^{\gamma}$ et, comme $\omega^{\gamma} \geqq \gamma$, on trouve $\omega^{\gamma} = \gamma$. Ainsi γ est un nombre epsilonien. Or, $\xi = \alpha\omega = \gamma\omega$.

Chaque solution $\xi > \omega$ de l'équation (1) est donc de la forme $\xi = \varepsilon\omega$, où ε est un nombre epsilonien. D'autre part, soit ε un nombre epsilonien quelconque. On a donc $\varepsilon = \omega^\varepsilon$ et, pour $\xi = \varepsilon\omega$ on trouve

$$\xi^\omega = (\varepsilon\omega)^\omega = (\omega^\varepsilon . \omega)^\omega = \omega^{(\varepsilon+1)\omega} = \omega^{\varepsilon\omega} = \omega^\xi,$$

et le nombre ξ satisfait à l'équation (1).

Nous avons ainsi démontré le

Théorème. *Pour qu'un nombre ordinal $\xi \neq \omega$ satisfasse à l'équation (1), il faut et il suffit qu'il soit de la forme $\xi = \varepsilon\omega$, où ε est un nombre epsilonien.*

Or, il est à remarquer que, quel que soit le nombre ordinal transfini α de seconde espèce, il existe un nombre ordinal $\xi > \alpha$, tel que

(2) $$\alpha^\xi = \xi^\alpha.$$

En effet, soit ε un nombre epsilonien $> \alpha$ et posons $\xi = \varepsilon\alpha$. Vu que $\alpha \geqq \omega$, nous aurons $\xi = \varepsilon\alpha > \varepsilon > \alpha$, d'où $\xi > \alpha$. Or, comme $\alpha < \varepsilon$, on a $\alpha^\varepsilon = \varepsilon$, d'où

$$\xi^\alpha = (\varepsilon\alpha)^\alpha = (\alpha^\varepsilon . \alpha)^\alpha = \alpha^{(\varepsilon+1)\alpha} = \alpha^{\varepsilon\alpha} = \alpha^\xi$$

(puisque, α étant un nombre de seconde espèce, on a $(\varepsilon+1)\alpha = \varepsilon\alpha$).

Le nombre ξ satisfait donc à l'équation (2).

Remarque. On définit les nombres epsiloniens comme nombres ordinaux ε satisfaisant à l'équation $\omega^\varepsilon = \varepsilon$. Or, on les pourrait définir aussi comme nombres ordinaux $\xi > \omega$ satisfaisant à l'équation

(3) $$2^\xi = \xi.$$

En effet, d'une part, si ε est un nombre ordinal tel que $\omega^\varepsilon = \varepsilon$, on a $\varepsilon \leq 2^\varepsilon \leq \omega^\varepsilon$, donc $2^\varepsilon = \varepsilon$. D'autre part, si $2^\xi = \xi$ et $\xi > \omega$, on a $\xi = \omega + \varrho$, où $\varrho \geqq 1$. Comme $2^\xi = \xi$, on trouve $\xi = 2^{2^\xi} = 2^{2^{\omega+\varrho}} = 2^{\omega . 2^\varrho} = \omega^{2^\varrho} \geqq \omega^2$, d'où $\omega + \xi = \xi = \omega + \varrho$, ce qui donne $\xi = \varrho$, donc, d'aprés (3) : $\xi = \omega^{2^\varrho} = \omega^{2^\xi}$, d'où $\xi = \omega^\xi$, ce qui prouve que ξ est un nombre epsilonien.

Quant à l'équation (3), outre les nombres epsiloniens, elle a encore la solution $\xi = \omega$.

(Reçu le 30 mai 1949)

Resolutions of the identity for commutative B^*-algebras of operators.

By NELSON DUNFORD in New Haven, Conn.

In this note we give a generalization of the well-known spectral formula $f(T)=\int f(\lambda)\,dE_\lambda$. The theorem we prove (Theorem 3) is based upon a theorem of GELFAND—NEUMARK[1]) and upon a generalization of F. RIESZ's[2]) classical result concerning the linear functionals on the space C. For the sake of completeness these two theorems (Theorems 1 and 2) are stated explicitly but without proof. The general idea of the proof was discovered by F. RIESZ[3]).

§ 1. Preliminary concepts.

1. 1. **The Riesz representation theorem.** In what follows we shall be concerned with a compact (= bicompact) Hausdorff space Λ and the family $C(\Lambda)$ of complex valued continuous functions defined on Λ. The space $C(\Lambda)$ is a complex Banach space under the norm $|f|=\sup\limits_{\lambda}|f(\lambda)|$ and as such determines a conjugate space $C^*(\Lambda)$, namely the space of all continuous linear maps of $C(\Lambda)$ into the complex number system. Besides being linear spaces $C(\Lambda)$ and $C^*(\Lambda)$ are partially ordered via the following definitions: for $f\in C(\Lambda)$ we say that $f\geqq 0$ if $f(\lambda)\geqq 0$, $\lambda\in\Lambda$; and for $x^*\in C^*(\Lambda)$ we say that $x^*\geqq 0$ if $x^*f\geqq 0$ for every $f\geqq 0$. It will be necessary in what follows to have a representation for the space $C^*(\Lambda)$. A representation for $C^*(\Lambda)$ was originally found, in the case where Λ is an interval of real numbers, by F. RIESZ. RIESZ's theorem has been successively generalized by RADON, SAKS, VON

[1]) I. GELFAND—M. NEUMARK, On the embedding of normed rings into the ring of operators in Hilbert space, *Mat. Sbornik*, N. S., **12** (1943), pp. 197—213. See also the very elegant proof of R. ARENS, On a theorem of Gelfand—Neumark, *Proceedings National Academy of Sciences*, **32** (1946), pp. 237—239.

[2]) F. RIESZ, Sur les opérations fonctionnelles linéaires, *Comptes Rendus de l'Académie des Sciences Paris*, **149** (1909), pp. 974—977.

[3]) F. RIESZ, Über quadratische Formen von unendlich vielen Veränderlichen, *Göttinger Nachrichten*, 1910, pp. 190—195.

Neumann, Kakutani and others. In order to state here the form of the generalized Riesz theorem that we shall need, it is necessary to introduce the notion of a regular measure on Λ. A function μ is called a *regular measure* on Λ if it is a countably additive complex valued set function defined on the field B of all Borel subsets of Λ and in case it also has the following property: for every Borel set $e \subset \Lambda$ and every $\varepsilon > 0$ there is a closed set $a \subset e$ and an open set $b \supset e$ such that for every $e' \in B$ with $a \subset e' \subset b$ we have $|\mu(e) - \mu(e')| < \varepsilon$. It is not difficult to see that the regular measures on Λ form a complex Banach space under the norm $|\mu| =$ total variation of $\mu(e)$, $e \in B$. For brevity we shall denote this complex Banach space of regular measures on Λ by the symbol $R(\Lambda)$. The space $R(\Lambda)$ is partially ordered by the definition $\mu \geqq 0$ if and only if $\mu(e) \geqq 0$, $e \in B$. With these notions in mind then the general form of the Riesz representation theorem may be stated as follows.

Theorem 1. *For every $x^* \in C^*(\Lambda)$ there is a uniquely determined point $\mu \in R(\Lambda)$ for which*

$$x^* f = \int_\Lambda f(\lambda)\, d\mu(\lambda), \qquad f \in C(\Lambda).$$

This correspondence $x^* \to \mu$ is an order preserving, isometric isomorphism between the spaces $C^*(\Lambda)$ and $R(\Lambda)$.

A proof of the theorem, in essentially this form, may be found in a paper of Kakutani[4]). Kakutani considers only the real continuous functions on Λ and represents the positive functionals. A proof of the theorem as we have stated it here may be based upon Kakutani's work and presents no real difficulty.

1. 2. **The Gelfand—Neumark representation theorem.** The particular compact Hausdorff space Λ to which Theorem 1 is to be applied is the space of maximal ideals of a certain commutative B^*-algebra. By a B-algebra we shall mean a complex Banach space $\mathfrak{A}$ with the properties

(α) There is a binary operation defined in $\mathfrak{A}$ with the properties

$$x(\lambda y + \mu z) = \lambda xy + \mu xz, \quad (\lambda y + \mu z)x = \lambda yx + \mu zx, \quad x(yz) = (xy)z,$$

where λ, μ are complex numbers and $x, y, z \in \mathfrak{A}$.

(β) There is a unit $e \in \mathfrak{A}$ with the property that $ex = x$, $x \in \mathfrak{A}$.

(γ) $|xy| \leqq |x| . |y|$ for $x, y \in \mathfrak{A}$; $|e| = 1$.

A B-algebra $\mathfrak{A}$ is called a B^*-algebra in case there is a unary operation $*$ in $\mathfrak{A}$ with the properties

(δ) $(x^*)^* = x$, $(\alpha x)^* = \bar{\alpha} x^*$, $(xy)^* = y^* x^*$, $(x+y)^* = x^* + y^*$, $|x^* x| = |x|^2$.

A B-algebra $\mathfrak{A}$ is called commutative in case $xy = yx$, $x, y \in \mathfrak{A}$. If λ is a

[4]) S. Kakutani, A concrete representation of abstract (M)-spaces, *Annals of Math.*, **42** (1941), pp. 994–1024; in particular Theorem 9, p. 1009.

maximal ideal in a commutative B-algebra, GELFAND[5]) has shown that the quotient algebra $\mathfrak{A}/\lambda$ is the complex number system. The complex number corresponding to an element $x \in \mathfrak{A}$ under the natural homomorphism $\mathfrak{A} \to \mathfrak{A}/\lambda$ will be denoted by $x(\lambda)$. If in the set Λ of all maximal ideals of $\mathfrak{A}$ we define neighborhoods of a point $\lambda_0 \in \Lambda$ to be all sets of the form

$$N(\lambda_0) = [\lambda:\ \lambda \in \Lambda,\ |x_i(\lambda) - x_i(\lambda_0)| < \varepsilon; i = 1, \ldots, n],$$

where x_i $(i = 1, \ldots, n)$ are arbitrary elements of $\mathfrak{A}$ and ε is an arbitrary positive number, then, as GELFAND has shown, Λ becomes a compact Hausdorff space. The functions $x(\lambda)$ are continuous on Λ and $\sup|x(\lambda)| \leq |x|$. In case the commutative B-algebra is also a B^*-algebra, much more can be said; namely,

T h e o r e m 2. (GELFAND—NEUMARK). *Let Λ be the compact Hausdorff space of maximal ideals of the commutative B^*-algebra $\mathfrak{A}$. Then $x^*(\lambda) = \overline{x(\lambda)}$ (the complex conjugate of $x(\lambda)$). Furthermore, the map $x \to x(\lambda)$ of $\mathfrak{A}$ into $C(\Lambda)$ is an isometric isomorphism of $\mathfrak{A}$ onto $C(\Lambda)$.*

§ 2. Resolutions of the identity for commutative B^*-algebras of operators in Hilbert spaces.

2. 1. **The algebras $\mathfrak{A}$ and $\mathfrak{A}(H)$.** Let H be a Hilbert space (not necessarily separable) and let $\mathfrak{A}(H)$ be the algebra of all continuous linear transformations in H. For $T \in \mathfrak{A}(H)$ let T^* be the adjoint of T and let $|T| = \sup\limits_{|x|=1} |Tx|$. Then $\mathfrak{A}(H)$ is a B^*-algebra with unit I. In what follows the symbol $\mathfrak{A}$ will be used for any closed commutative B^* subalgebra of $\mathfrak{A}(H)$ which contains the unit I. The symbol Λ will be used for the compact Hausdorff space of maximal ideals in $\mathfrak{A}$. For a function $f \in C(\Lambda)$ the symbol $\bar{f}$ will denote the function $\bar{f}(\lambda) = \overline{f(\lambda)}$. Thus, according to Theorem 2, there is an isometric isomorphism, $f \leftrightarrow T(f)$, between the algebras $C(\Lambda)$ and $\mathfrak{A}$, and this isomorphism has the property that $T(f)^* = T(\bar{f})$.

2. 2. **The resolution of the identity corresponding to the isomorphism $\mathbf{f \leftrightarrow T(f)}$.** Let f, $T(f)$ be corresponding elements of $C(\Lambda)$ and $\mathfrak{A}$ under some isometric isomorphism. The symbol B will stand for the Borel subsets of Λ.

T h e o r e m 3. *For $e \in B$ there is a uniquely determined $E_e \in \mathfrak{A}(H)$ with the properties*

(i) *For $x \in H$ the function $E_e x$ is countably additive on B.*
(ii) *For every pair $x, y \in H$, the scalar product $(E_e x, y)$ is a regular countably additive set function on B whose total variation is at most $|x|\,|y|$.*

---

[5]) I. GELFAND, Normierte Ringe, *Mat. Sbornik*, N. S., **9** (1941), pp. 1—24.

(iii) *For every* $T \in \mathfrak{A}$ *and* $e, e_1, e_2 \in B$

$$E_e T = T E_e, \ E_{e_1} E_{e_2} = E_{e_2} E_{e_1}, \ E_e^2 = E_e, \ E_e^* = E_e.$$

(iv) $(T(f)x, y) = \int_\Lambda f(\lambda) d(E_\lambda x, y), \ x, y \in H, \ f \in C(\Lambda).$ [6])

The number $(T(f)x, y)$ is clearly linear in f and since $|(T(f)x, y)| \leq \leq |T(f)||x||y| = |f||x||y|$ there is, by Theorem 1, a uniquely determined regular measure $\mu(e, x, y)$ on Λ with

(a) $$(T(f)x, y) = \int_\Lambda f(\lambda)\, d\mu(\lambda, x, y), \quad f \in C(\Lambda),$$

(b) $$|\mu(e, x, y| \leq \operatorname*{var}_{e \in B} \mu(e, x, y) \leq |x||y|, \quad e \in B.$$

Since $(T(f)\alpha x, y) = \alpha(T(f)x, y)$ we have from (a) $\int_\Lambda f(\lambda)\, d\mu(\lambda, \alpha x, y) = = \int_\Lambda f(\lambda)\, d\alpha\mu(\lambda, x, y)$ and since the regular measure is uniquely determined by the functional it is seen that $\mu(e, \alpha x, y) = \alpha\mu(e, x, y)$. Similarly it may be shown that $\mu(e, x, y)$ is bilinear in x, y. It is also Hermitian symmetric for if f is real the operator $T(f) = T(\bar{f}) = T(f)^*$ is self-adjoint and $[T(f)x, y] = = \overline{(T(f)y, x)}$. Thus by (a), $\int_\Lambda f(\lambda)\, d\mu(\lambda, x, y) = \int_\Lambda f(\lambda)\, \overline{d\mu(\lambda, y, x)}$ for f real, and again by the uniqueness argument $\mu(e, x, y) = \overline{\mu(e, y, x)}$. Hence $\mu(e, x, y)$ is an Hermitean symmetric bilinear form of norm (see (b)) at most unity. It follows from the well-known elementary theorem which gives the representation of such forms that there is a self-adjoint operator E_e of norm $|E_e| \leq 1$ with $\mu(e, x, y) = (E_e x, y)$. This fact when combined with (a) and (b) proves (ii) and (iv). To see that $TE_e = E_e T$ let f, g be arbitrary elements of $C(\Lambda)$ and since $T(f)$, $T(g)$ commute we have $(T(f)\,T(g)x, y) = (T(f)x, \ T(g)^* y)$. Equation (iv) then gives $\int f(\lambda) d(E_\lambda T(g)x, y) = \int f(\lambda)\, d(E_\lambda x, T(g)^* y)$ and since the regular measure is uniquely determined by the integral,

$$(E_e T(g)x, y) = (E_e x, T(g)^* y) = (T(g)E_e x, y), \text{ i. e., } E_e T(g) = T(g)E_e.$$

We now proceed to a proof of the relation $E_{e_1} E_{e_2} = E_{e_1 e_2}$. For this we shall need the following lemmas.

Lemma 1. *If* $f(\lambda) \geq 0$ *is continuous and* $\mu(e)$ *is a non-negative regular measure on* Λ *and if* C *is a closed set in* Λ *then*

$$\text{g. l. b.} \int_\Lambda f(\lambda)\, g(\lambda)\, d\mu(\lambda) = \int_C f(\lambda)\, d\mu(\lambda),$$

[6]) If the algebra is separable, or if the Hilbert space itself is separable, the theorem is a corollary of known results. See J. v. Neumann, Zur Algebra der Funktionaloperationen und Theorie der normalen Operatoren, *Math. Annalen*, **102** (1929), pp. 370–427, Satz 10, or Béla v. Sz. Nagy, *Spektraldarstellung linearer Transformationen des Hilbertschen Raumes* (Berlin, 1942), pp. 66–67.

where the g. l. b. is taken over $g \in C(\Lambda)$ for which $g(\lambda) \geqq \varphi_C(\lambda)$, the characteristic function of C.

We shall first prove the lemma in case $f(\lambda) \equiv 1$. Let G be an open set containing C. Since Λ is a normal topological space, there is a $g \in C(\Lambda)$ with

$$0 \leqq g(\lambda) \leqq 1; \quad g(\lambda) = 1 \text{ for } \lambda \in C; \quad g(\lambda) = 0 \text{ for } \lambda \bar{\in} G.$$

Thus

$$\mu(G) \geqq \int_\Lambda g(\lambda)\, d\mu(\lambda) \geqq \mu(C)$$

and since μ is regular, $\mu(G)$ differs from $\mu(C)$ by as little as we please. This proves the lemma in case $f(\lambda) \equiv 1$. For the general case we let $\mu_1(e) = \int_e f(\lambda)\, d\mu(\lambda)$ and observe that μ_1 is a non-negative regular measure which satisfies the equation

$$\text{(c)} \qquad \int_\Lambda g(\lambda) f(\lambda)\, d\mu(\lambda) = \int_\Lambda g(\lambda)\, d\mu_1(\lambda)$$

providing g is the characteristic function of a Borel set in Λ. Since an arbitrary $g \in C(\Lambda)$ may be uniformly approximated by linear combinations of characteristic functions of Borel sets, it is seen that (c) holds for every continuous function. The desired conclusion follows then by an application of the case $f(\lambda) \equiv 1$, already proved, to equation (c).

Lemma 2. $(E_e x, x) \geqq 0$ *for* $e \in B$, $x \in H$.

In view of Theorem 1 it will suffice to show that $\int f(\lambda) d(E_\lambda x, x) \geqq 0$ for every continuous positive function f. For such an f, $f^{\frac{1}{2}}(\lambda)$ is real and hence $T(f^{\frac{1}{2}})$ is self-adjoint. Thus $0 \leqq |T(f^{\frac{1}{2}}) x|^2 = (T(f^{\frac{1}{2}}) x, T(f^{\frac{1}{2}}) x) =$
$= (T(f) x, x) = \int_\Lambda f(\lambda) d(E_\lambda x, x)$, q. e. d.

Now let f, g be non-negative and continuous. Then by Lemma 2 we have

$$0 \leqq \int f(\lambda) g(\lambda)\, d(E_\lambda x, x) = (T(fg) x, x) = (T(f) x, T(g) x) = \\ = \int f(\lambda)\, d(E_\lambda x, T(g) x).$$

Since $(E_e x, T(g) x) = (T(g) x, E_e x) = \int g(\mu)\, d(E_\mu x, E_e x)$, the above equation may be written as

$$\text{(d)} \qquad 0 \leqq \int f(\lambda) g(\lambda)\, d(E_\lambda x, x) = \int f(\lambda)\, d \int g(\mu)\, d(E_\mu x, E_\lambda x).$$

Since f and g are arbitrary non-negative continuous functions, we see from Theorem 1 first that $\int g(\mu)\, d(E_\mu x, E_e x) \geqq 0$ and then that

$$\text{(e)} \qquad (E_{e_1} x, E_e x) \geqq 0 \text{ for } e, e_1 \in B.$$

Let C be a closed set, let φ be the characteristic function of C and let $\mu(e)=(E_{eC}x, x)$. Then from (e)

$$\int g(\mu)\, d(E_\mu x, E_e x) \geqq (E_C x, E_e x), \quad g \geqq \varphi,$$

and so by (d)

$$\int f(\lambda)\, g(\lambda)\, d(E_\lambda x, x) \geqq \int f(\lambda)\, d(E_C x, E_\lambda x); \quad f \geqq 0, \ g \geqq \varphi.$$

By Lemma 1, then

$$\int f(\lambda)\, d\mu(\lambda) = \int_C f(\lambda)\, d(E_\lambda x, x) \geqq \int f(\lambda)\, d(E_C x, E_\lambda x)$$

which gives

$$\text{(f)} \qquad (E_{eC}x, x) = \mu(e) \geqq (E_C x, E_e x) = (E_e E_C x, x).$$

Now let $g \geqq \varphi$, then by (d)

$$\int f(\lambda)\, d\mu(\lambda) = \int_C f(\lambda)\, d(E_\lambda x, x) \leqq \int f(\lambda)\, g(\lambda)\, d(E_\lambda x, x) =$$

$$= \int f(\lambda)\, d \int g(\mu)\, d(E_\mu x, E_\lambda x),$$

and since the measures are regular,

$$\mu(e) = (E_{eC}x, x) \leqq \int g(\mu)\, d(E_\mu x, E_e x); \ e \in B.$$

An application of Lemma 1 gives $(E_{eC}x, x) \leqq (E_C x, E_e x) = (E_e E_C x, x)$, which when combined with (f) yields $(E_{eC}x, x) = (E_e E_C x, x)$. Assume for the moment that E_e and E_C commute, then the operator $A = E_{eC} - E_e E_C$ is self-adjoint and since $(Ax, x) = 0$ it follows that $A = 0$. For fixed $e \in B$ the two regular measures $(E_{ea}x, y)$, $(E_e E_a x, y)$ coincide for closed sets a. They must therefore coincide everywhere on B. Thus $E_e E_a = E_{ea}$, $E_e^2 = E_e$. It remains to be shown that $E_{e_1} E_e = E_e E_{e_1}$. But this follows immediately from the fact that E_e is self-adjoint and commutes with $T(f)$. For from the equation $(T(f) E_{e_1} x, y) = (T(f) x, E_{e_1} y)$ we get $\int f(\lambda)\, d(E_\lambda E_{e_1} x, y) = \int f(\lambda)\, d(E_\lambda x, E_{e_1} y)$ and thus $(E_e E_{e_1} x, y) = (E_e x, E_{e_1} y) = (E_{e_1} E_e x, y)$. Only statement (1) remains to be proved. Let e_n be disjoint, $e_n \in B$, $e = \bigcup_n e_n$ and $a_n = e - \bigcup_{m=1}^{n} e_m$. Then we must show that $E_{a_n} x \to 0$. But $|E_{a_n} x|^2 = (E_{a_n} x, E_{a_n} x) = (E_{a_n} x, x) \to 0$ by (ii). This completes the proof of the theorem.

(Received August 16, 1949.)

Die Reziprozitätsformel für Dedekindsche Summen.

Von H. RADEMACHER in Philadelphia.

Ich erinnere mich gern an einen Besuch in Budapest vor 20 Jahren, bei welcher Gelegenheit ich einen Vortrag vor einer Gruppe von FEJÉRS Freunden und Schülern halten durfte[1]). Die jetzige Gelegenheit gibt mir einen Anlaß, auf meinen damaligen Gegenstand, einen Reziprozitätssatz über gewisse Summen, die ich kurz Dedekindsche Summen nennen möchte, zurückzukommen.

1. Eine Dedekindsche Summe $s(h, k)$ ist folgendermaßen definiert. Es sei zur Abkürzung für reelle x

$$\text{(1a)}\qquad \begin{aligned} ((x)) &= x-[x]-1/2 && \text{für nicht-ganze } x, \\ ((x)) &= 0 && \text{für ganze } x \end{aligned}$$

gesetzt. Dann sei für zueinander teilerfremde h und k

$$\text{(1b)}\qquad s(h, k)=\sum_{\mu=1}^{k-1}\left(\left(\frac{\mu}{k}\right)\right)\left(\left(\frac{h\mu}{k}\right)\right).$$

Die erwähnte *Reziprozitätsformel* lautet nun

$$\text{(2)}\qquad s(h, k)+s(k, h)=-\frac{1}{4}+\frac{h}{12k}+\frac{k}{12h}+\frac{1}{12hk}.$$

Die Dedekindschen Summen erscheinen zuerst in Dedekinds Kommentar zu Riemanns nachgelassenen Fragmenten über die Modulfunktionen[2]). Dort findet man auch Dedekinds Beweis für die Reziprozitätsformel, der auf der Gruppeneigenschaft der Modulsubstitutionen beruht. Die Dedekindschen Summen, die eine fundamentale Rolle in der Theorie der Theta- und Modulfunktionen spielen, treten auch in der von HARDY und RAMANUJAN geschaffenen Theorie der Partitionen auf. Die Reziprozitätsformel (2) ist jedoch eine rein zahlentheoretische Aussage. Ich habe dafür mehrere von der Theorie der Modulfunktionen unabhängige Beweise gegeben, zu denen ich hier einen besonders durchsichtigen hinzufügen möchte.

---

[1]) H. RADEMACHER, Egy reciprocitásképletről a modulfüggvények elméletéből, *Matematikai és Fizikai Lapok*, **40** (1933), S. 24—34.

[2]) Erläuterungen zu Riemanns Fragmenten über die Grenzfälle der elliptischen Modulfunktionen, *Riemanns Werke* (1876), S. 438—447, *Dedekinds Werke* (1930), Bd. 1, S. 159—172.

2. Wir benötigen zwei Lemmata:

Lemma I. *Sind m und n zwei teilerfremde positive ganze Zahlen, so ist, mit den Bezeichnungen* (1a),

$$\int_0^1 ((mx))\,((nx))\,dx = \frac{1}{12\,mn}.\text{[3]} \tag{3}$$

Lemma II. *Es seien $f(x)$, $g(x)$, $h(x)$ Funktionen von beschränkter Schwankung im Intervall $a \leq x \leq b$, die paarweise keine Unstetigkeit gemeinsam haben. Dann gilt die Gleichung*

$$\int_a^b f(x)\,dg(x)h(x) = \int_a^b f(x)g(x)\,dh(x) + \int_a^b f(x)h(x)\,dg(x) \tag{4}$$

zwischen Stieltjes-Integralen.

Bemerkung. Für $f(x)$ identisch gleich 1 geht (4) über in die Formel für partielle Integration

$$\Big[g(x)h(x)\Big]_a^b = \int_a^b g(x)\,dh(x) + \int_a^b h(x)\,dg(x) \tag{5}$$

unter den obigen Bedingungen für $g(x)$, $h(x)$.

Was den *Beweis* von Lemma II angeht, so bemerken wir zunächst, daß $g(x)h(x)$ von beschränkter Schwankung ist und daß dieses Produkt keine Unstetigkeit mit $f(x)$ gemeinsam hat.

Wir wissen ferner, daß

$$\int_a^b f(x)\,d\varphi(x)$$

existiert, wenn $f(x)$ und $\varphi(x)$ beide von beschränkter Schwankung sind und keine Unstetigkeit gemeinsam haben[4]). Die in (4) genannten Integrale existieren also alle.

Haben wir nun eine Einteilung des Intervalls (a, b)

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$$

vor uns und haben $x_{j-1} \leq \xi_j \leq x_j$ gewählt, so betrachten wir die Identität

$$\begin{aligned}\sum_{j=1}^{n} f(\xi_j)\{g(x_j)h(x_j) - g(x_{j-1})h(x_{j-1})\} &= \\ = \sum_{j=1}^{n} f(\xi_j)g(x_j)\,\{h(x_j) - h(x_{j-1})\} &+ \sum_{j=1}^{n} f(\xi_j)h(x_{j-1})\{g(x_j) - g(x_{j-1})\}.\end{aligned} \tag{6}$$

Hier verlangen nur die Summen auf der rechten Seite einige Überlegung. Wenn man $f(x)$, $g(x)$, $h(x)$ als Differenzen von positiven, monoton wachsen-

---

[3]) Einen Beweis findet man in LANDAU's „Vorlesungen über Zahlentheorie", Bd. 2, S. 171; siehe auch J. FRANEL, *Göttinger Nachrichten* 1924, S. 199 und E. LANDAU, *ibidem*, S. 203.

[4]) Vgl. z. B. D. V. WIDDER, *The Laplace Transform* (Princeton, 1941), p. 25, Theorem 14.

den Funktionen schreibt, kommt man auf Summen von der Form

(7a) $$\sum \varphi(\xi_j)\,\psi(x_j)\{\chi(x_j)-\chi(x_{j-1})\}$$

und

(7b) $$\sum \varphi(\xi_j)\,\psi(x_{j-1})\{\chi(x_j)-\chi(x_{j-1})\},$$

wo φ, ψ, χ positive und monoton wachsende Funktionen sind, die paarweise keine Unstetigkeiten gemeinsam haben. Diese Summen liegen nun offenbar zwischen

$$\sum \varphi(x_{j-1})\,\psi(x_{j-1})\{\chi(x_j)-\chi(x_{j-1})\}$$

und

$$\sum \varphi(x_j)\,\psi(x_j)\{\chi(x_j)-\chi(x_{j-1})\},$$

die beide nach dem oben zitierten Satz mit Verfeinerung der Teilung gegen

$$\int_a^b \varphi(x)\,\psi(x)\,d\chi(x)$$

konvergieren, gegen welchen Limes daher auch (7a) und (7b) streben. Somit folgt aus (6) die zu beweisende Gleichung (4).

3. Unser Beweis der Reziprozitätsformel (2) ist nun schnell geführt, indem wir das Stieltjes-Integral

(8) $$I_\varepsilon = \int_\varepsilon^{1-\varepsilon} ((x))\,d((hx))((kx)),\quad \varepsilon > 0$$

auf zwei verschiedene Weisen behandeln. Die Funktion $((x))$ ist stetig im Intervall $\varepsilon \leqq x \leqq 1-\varepsilon$, und die Funktionen $((hx))$ und $((kx))$ haben dort wegen $(h, k) = 1$ keine Unstetigkeiten gemeinsam.

Einerseits ist nun wegen (4)

$$I_\varepsilon = \int_\varepsilon^{1-\varepsilon} ((x))((hx))\,d((kx)) + \int_\varepsilon^{1-\varepsilon} ((x))((kx))\,d((hx)) =$$

$$= \int_\varepsilon^{1-\varepsilon} ((x))((hx))\,k\,dx - \sum_{\varepsilon<\frac{\mu}{k}<1-\varepsilon} \left(\left(\frac{\mu}{k}\right)\right)\left(\left(\frac{h\mu}{k}\right)\right) + \int_\varepsilon^{1-\varepsilon} ((x))((kx))\,h\,dx - \sum_{\varepsilon<\frac{\nu}{h}<1-\varepsilon} \left(\left(\frac{\nu}{h}\right)\right)\left(\left(\frac{k\nu}{h}\right)\right)$$

und daher

$$\lim_{\varepsilon\to+0} I_\varepsilon = k\int_0^1 ((x))((hx))\,dx - \sum_{\mu=1}^{k-1} \left(\left(\frac{\mu}{k}\right)\right)\left(\left(\frac{h\mu}{k}\right)\right) + h\int_0^1 ((x))((kx))\,dx - \sum_{\nu=1}^{h-1} \left(\left(\frac{\nu}{h}\right)\right)\left(\left(\frac{k\nu}{h}\right)\right).$$

Infolge von (3) und (1b) ergibt sich somit

(9) $$\lim_{\varepsilon\to+0} I_\varepsilon = \frac{k}{12h} - s(h, k) + \frac{h}{12k} - s(k, h).$$

Anderseits erhält man mittels der partiellen Integration (5) aus (8)

$$I_\varepsilon = \Big[((x))((hx))((kx))\Big]_\varepsilon^{1-\varepsilon} - \int_\varepsilon^{1-\varepsilon} ((hx))((kx))\,d((x)) =$$

$$= -2((\varepsilon))((h\varepsilon))((k\varepsilon)) - \int_\varepsilon^{1-\varepsilon} ((hx))((kx))\,dx$$

und somit

$$\lim_{\varepsilon \to +0} I_\varepsilon = -2\{\lim_{\varepsilon \to +0} ((\varepsilon))\}^3 - \int_0^1 ((hx))((kx))\,dx$$

oder, infolge von (1a) und (3),

$$\lim_{\varepsilon \to +0} I_\varepsilon = \frac{1}{4} - \frac{1}{12hk}. \tag{10}$$

Die Gleichungen (9) und (10) ergeben nun zusammen die zu beweisende Formel (2).

(Eingegangen am 8. August 1949.)

The location of critical points of harmonic functions.

By J L. WALSH in Cambridge, Mass.

There has recently been developed an extensive theory of the location of the critical points of harmonic functions, namely the study of the geometry of regions containing some or all critical points of a given harmonic function; compare a forthcoming volume by the present writer in the Colloquium Series of the American Mathematical Society. The most useful method in establishing that theory is the interpretation of the given harmonic function as the potential due to a suitable distribution of matter, and detailed study of the corresponding field of force. The latter is usually set up by means of GREEN's theorem or CAUCHY's integral. It is to be expected that the classical results of F. RIESZ on the representation of superharmonic functions [1]) would also be of significance here; indeed, it is the purpose of the present note to indicate the great power of his results in this study of critical points when superharmonic functions are involved.

In the plane of $z=x+iy$, a function $u(z)\equiv u(x,y)$ is *harmonic* at a finite point z if throughout a neighborhood of that point the function is continuous together with its first and second partial derivatives, and satisfies LAPLACE's equation. A function $u(z)$ is harmonic *in a region* if it is harmonic at every point of that region. A finite point (x_0, y_0) is a *critical point* of a function $u(x,y)$ harmonic at (x_0, y_0) if the two first partial derivatives of $u(x,y)$ vanish there. A critical point remains a critical point under one-to-one conformal transformation.

We state for reference only a small part of the results of F. RIESZ:

If the function $u(z)$ is superharmonic in the (regular) region R, and if there exists a function harmonic in R inferior to $u(z)$ throughout R, then $u(z)$ can be represented in R in the form

$$(1) \qquad u(z)\equiv\int_R g(z,t,R)\,d\mu(e)+h(z), \quad d\mu(e)\geqq 0,$$

where $g(z,t,R)$ is Green's function for R with pole in z, where as t varies

---

1) F. RIESZ, Sur les fonctions subharmoniques et leur rapport à la théorie du potentiel. II, *Acta Math.*, **54** (1930), pp. 321–360.

over R the integral is taken with respect to a suitably chosen positive additive set function $\mu(e)$ defined for t on all open sets e whose closures lie in R, and where $h(z)$ is harmonic in R.

In the particular region R_1: $|z|<1$ we shall consider the NE lines, that is to say, the lines of non-euclidean (hyperbolic) geometry, namely the arcs in R_1 of circles and straight lines orthogonal to the circumference $|z|=1$. We define NE lines in an arbitrary region that can be mapped one-to-one and conformally onto R_1 as the images of the NE lines in R_1, and shall establish

Theorem 1. *Let R be the interior of a Jordan curve C, and let B be a closed set in R which together with C bounds a subregion R' of R. Let $u(z)$ be harmonic in R', superharmonic in R, continuous in the closed neighborhood of C in $R+C$, and zero on C. Then all critical points of $u(z)$ in R' lie in the smallest NE convex set Π of R containing B.*

We map R onto the interior of the unit circle $|z|=1$ so that an arbitrary preassigned point of R' not in Π is transformed into the origin 0. We preserve the original notation, and choose an arbitrary NE line through 0 wholly in R' as the axis of reals, with B in the upper half-plane. It remains to show merely that 0 is not a critical point of $u(z)$. Equation (1) is a consequence of Riesz's theorem. Since both $u(z)$ and $h(z)$ are harmonic in R', it follows from (1) that the integral is harmonic in R', and follows further that we have $d\mu=0$ in R'. Consequently the integral may be taken not over R but over an arbitrary open set in R containing $R-R'$. The function $g(z,t,R)$ considered as a function of z for t in $R-R'$ can be extended harmonically by reflection across C; the values at points mutually inverse with respect to C are the negatives of each other, so the integral represents a function of z harmonic in an annular region containing C, a function vanishing on C itself. It follows from (1) that $h(z)$ is continuous on $R+C$ and zero on C, hence identically zero in R. Thus equation (1) reduces to

$$u(z)=\int_{R-R'} g(z,t,R)\,d\mu(e), \quad z \text{ in } R; \tag{2}$$

more accurately, the integral is to be taken over an arbitrary open subset of R containing $R-R'$.

We can write explicitly for z in R

$$g(z,t,R)\equiv\log|z-1/\bar{t}|-\log|z-t|+\log|t|,$$

so if we set $f(z)\equiv u(z)+iv(z)$, where $v(z)$ is conjugate to $u(z)$ in R', we have for z in R for a suitable determination of the multiple-valued functions involved,

$$f(z)=\int_{R-R'}[\log(z-1/\bar{t})-\log(z-t)+\log|t|]\,d\mu+\text{const},$$

$$f'(z) = \int\limits_{R-R'} \frac{d\mu}{z - 1/\bar{t}} - \int\limits_{R-R'} \frac{d\mu}{z-t}, \quad z \text{ in } R'; \tag{3}$$

of course $f'(z)$ is single-valued in R'.

The conjugate of the number $\mu/(z-\alpha)$ where μ is real, represents generically the force at the point z due to a particle of mass μ at the point α, where the particle repels with a force equal to the mass divided by the distance. Thus the conjugate of the second member of (3) represents the force at z due to a spread of positive mass at the (finite) points $1/\bar{t}$ and numerically equal negative mass at the points t. The force at the particular point $z=0$ due to the mass μ (>0) at $1/\bar{t}$ and the mass $-\mu$ at t is

$$\frac{\mu}{-1/t} - \frac{\mu}{-\bar{t}} = \mu\left(\frac{1}{\bar{t}} - t\right),$$

which when t lies in R in the upper half-plane is a vector with a non-vanishing component vertically upward. Thus it follows from (3) that the total force at $z=0$ has a non-vanishing component vertically upward, so we have $f'(0) \neq 0$ and hence 0 is not a critical point of $u(z)$; Theorem 1 is established.

As a special case of Theorem 1 we have

Theorem 2. *Let R be the interior of a Jordan curve C, and let B be a closed set in R which together with C bounds a subregion R' of R. Let $u(z)$ be harmonic in R', continuous in $R'+B+C$, zero on C, and unity on B. Then all critical points of $u(z)$ in R' lie in the smallest NE convex set of R containing B.*

Extend the definition of $u(z)$ if necessary so that $u(z)$ is defined and equal to unity at every point of $R-(R'+B)$. Then $u(z)$ as thus extended is superharmonic in R, for $u(z)$ is locally superharmonic; it is sufficient to note that $u(z)$ is continuous in R, is harmonic at every point of R' and at every (interior) point of $R-(R'+B)$, and at an arbitrary point P of B is not less than the average of the values of $u(z)$ over any sufficiently small circumference whose center is P. Theorem 2 follows.

The contrast between Theorems 1 and 2 indicates the advantage of the methods of F. Riesz, for Theorem 2 can be proved[2]) without the use of (1) and (2) by establishing an equation similar to (3) where the integral is a line integral taken over an auxiliary variable point set B', chosen as a level locus of $u(z)$ in R', and consisting of mutually exterior analytic Jordan curves. Theorem 1 obviously applies to functions $u(z)$ much more general than those of Theorem 2, indeed applies to any linear combination with

[2]) J. L. Walsh, Note on the location of zeros of the derivative of a rational function whose zeros and poles are symmetric in a circle, *Bulletin American Math. Society,* **45** (1939), pp. 462–470.

positive constant coefficients of functions of the type of Theorem 2, each such function defined as above throughout R.

In Theorem 2 we may choose $u(z)$ a constant multiple of $-\log|R(z)|$, with

$$R(z)=\prod_{1}^{n}\frac{z-\alpha_k}{1-\bar{\alpha}_k z}, \qquad |\alpha_k|<1,$$

and where B is a level locus of $u(z)$; the critical points of $u(z)$ are precisely the critical points of $R(z)$, with the omission of the multiple zeros of $R(z)$. Theorem 2 then implies that *all critical points of $R(z)$ in R lie in the smallest NE convex set in R containing the points α_k;* this theorem is a NE analogue of the classical theorem of LUCAS regarding the zeros of the derivative of a polynomial; well-known beautiful applications of LUCAS's theorem have been made by L. FEJÉR.

As a further illustration of the method of F. RIESZ we prove

Theorem 3. *Let Π_1 and Π_2 be the upper and lower half-planes, let R be the region $|z|<1$, and let a subregion R' of R be bounded by the unit circle C and a closed set B in R not intersecting the axis of reals A. Let the function $u(x,y)$ be harmonic in R', continuous in the closed neighborhood of C in $R+C$, zero on C, superharmonic in $R\cdot\Pi_1$, and subharmonic in $R\cdot\Pi_2$. Then 0 is not a critical point of $u(x,y)$.*

We assume, as we may do with no loss of generality, that R' is symmetric in A and that we have $u(x,y)\equiv-u(x,-y)$ in R'; for we need merely replace a given $u(x,y)$ by the function $u_1(x,y)\equiv u(x,y)-u(x,-y)$ to obtain these conditions. Moreover we shall prove $\partial u_1(0,0)/\partial y\neq 0$, which implies $\frac{1}{2}\partial u_1(0,0)/\partial y=\partial u(0,0)/\partial y\neq 0$. We revert to the original notation, and note in particular that the condition $u(x,y)\equiv-u(x,-y)$ implies $u(x,0)\equiv 0$.

The function $u(z)$ is superior to the function zero in the region $R\cdot\Pi_1$, so by F. RIESZ's theorem we have for z in $R\cdot\Pi_1$

$$u(z)=\int_{R\cdot\Pi_1}g(z,t,R\cdot\Pi_1)\,d\mu+h_1(z), \quad d\mu\geqq 0, \tag{4}$$

where $h_1(z)$ is harmonic in $R\cdot\Pi_1$. Throughout the region $R'\cdot\Pi_1$ the functions $u(z)$ and $h_1(z)$ are harmonic and $d\mu$ is zero, so the integral can be taken over an arbitrary open set in $R\cdot\Pi_1$ containing the set $(R-R')\cdot\Pi_1$. The function $g(z,t,R\cdot\Pi_1)$ considered as a function of z can be extended harmonically across A, then across C, so the integral in (4) represents a function continuous on A and C and vanishing there. It follows that $h_1(z)$ is continuous in the closure of $R\cdot\Pi_1$ on the boundary of $R\cdot\Pi_1$, and zero on the boundary of $R\cdot\Pi_1$, hence identically zero in $R\cdot\Pi_1$.

From the definitions and uniqueness of the functions involved we verify for t in $(R-R')\cdot\Pi_1$ and z in R the identity $g(z,t,R)-g(z,\bar{t},R)\equiv g(z,t,R\cdot\Pi_1)$, where the latter function is defined in $R\cdot\Pi_2$ by harmonic extension. Consequently equation (4) for z in $R \cdot \Pi_1$ can be extended so as to be valid for arbitrary z in R:

$$u(z) \equiv \int_{(R-R') . \Pi_1} g(z, t, R) d\mu - \int_{(R-R') . \Pi_2} g(z, t, R) d\mu,$$

where the integrals are taken over arbitrary disjoint open sets in R containing respectively $(R-R') \cdot \Pi_1$ and $(R-R') \cdot \Pi_2$. If as in (3) we now set $f(z) \equiv u(z) + iv(z)$, where $v(z)$ is conjugate to $u(z)$ in R', the conjugate of the vector $f'(z)$ represents the force at a point z of R' due to a negative distribution $-\mu$ on $(R-R') \cdot \Pi_1$ and an equal positive distribution on the inverse of $(R-R') \cdot \Pi_1$ with respect to C, plus the force at z due to the positive distribution μ on $(R-R') \cdot \Pi_2$ and an equal negative distribution on the inverse of $(R-R') \cdot \Pi_2$ in C. All these distributions are located on bounded point sets. Consequently the total force at 0 has a non-zero component vertically upward, as has the conjugate of $f'(z) = \partial u/\partial x + i\partial v/\partial x = \partial u/\partial x - i\partial u/\partial y$ at 0, and the theorem follows.

In Theorem 3, obviously no point of A in R can be a critical point of $u(z)$.

Theorem 3 can be used in proving

Theorem 4. *Let R be the interior of a Jordan curve C, and let B_1 and B_2 be two disjoint closed sets in R which together form the boundary of a subregion R' of R. Let $u(z)$ be harmonic in R', continuous in $R' + B_1 + B_2 + C$, equal to zero on C, to unity on B_1, and to $-c(<0)$ on B_2. If a NE line λ for R separates B_1 and B_2, then no critical point of $u(z)$ lies on λ in R'. Consequently, if such lines λ exist, all critical points of $u(z)$ in R' lie in two NE convex closed subregions of R containing B_1 and B_2 respectively and separated by each NE line λ which separates B_1 and B_2.*

Map R onto the interior of the unit circle so that a given NE line λ is transformed into the axis of reals A and an arbitrarily chosen point of λ is transformed into the origin 0; we take B_1 in Π_1 and B_2 in Π_2, and we retain the original notation. It is sufficient to show that 0 is not a critical point. We define $u(z)$ as unity in the points of R separated from 0 by B_1, and as $-c$ in the points of R separated from 0 by B_2. Thus $u(z)$ is superharmonic in $R \cdot \Pi_1$ and subharmonic in $R \cdot \Pi_2$. Theorem 4 follows from Theorem 3.

Theorem 4 can be proved without the use of (4), but again the contrast between Theorems 3 and 4 indicates the power of the methods of RIESZ. Theorems 1 and 3 are presented here not to exhaust the range of the method but merely to illustrate the simplifying and unifying influence of the results due to F. RIESZ, by enabling us under broad conditions to set up and study a field of force corresponding to given superharmonic and subharmonic functions.

(Received August 22, 1949.)

A remark on functions of several complex variables.

By A. ZYGMUND in Chicago.

1. Let $f(z)$ be a function of a complex variable z, regular for $|z|<1$. The very well known result of NEVALINNA and OSTROWSKI asserts that, if

$$\int_0^{2\pi} \log^+ |f(re^{i\theta})|\, d\theta = O(1), \tag{1}$$

then $f(z)$ has a finite non-tangential limit at almost every point of the circumference $|z|=1$. Condition (1) is equivalent to the fact that the subharmonic and non-negative function $\log^+|f(z)|$ has a harmonic majorant for $|z|<1$. Since the existence of a harmonic majorant is an invariant of conformal mapping, the Nevanlinna—Ostrowski result may be stated for any domain limited by a simple and rectifiable curve, and even in a much more general case.

The situation is different for functions $f(z_1, z_2, \ldots, z_k)$ of several complex variables. Only in exceptional cases are two topologically equivalent domains in the space Z_k of the complex variables $z_1 = x_1 + iy_1, \ldots, z_k = x_k + iy_k$ equivalent through complex analytic mapping. Two simplest examples of such non-equivalent domains are the unit hypersphere

$$|z_1|^2 + |z_2|^2 + \ldots + |z_k|^2 < 1 \tag{S}$$

and the unit polycylinder

$$|z_1| < 1, |z_2| < 1, \ldots, |z_k| < 1. \tag{C}$$

As regards the latter, it has been shown (see [6]) that if $f(z_1, z_2, \ldots, z_k)$ is regular in C and if

$$\int_0^{2\pi} \ldots \int_0^{2\pi} \log^+ |f| (\log^+ \log^+ |f|)^{k-1} d\theta_1 \ldots d\theta_k = O(1) \qquad (0 \leqq r_1, \ldots, r_k < 1), \tag{2}$$

where f stands for $f(r_1 e^{i\theta_1}, r_2 e^{i\theta_2}, \ldots, r_k e^{i\theta_k})$, then f has a finite limit for almost every $(\theta_1^0, \ldots, \theta_k^0)$ as the point $(r_1 e^{i\theta_1}, \ldots, r_k e^{i\theta_k})$ approches $(e^{i\theta_1^0}, \ldots, e^{i\theta_k^0})$ along any non-tangential path. Though the problem is open, there seems to be little doubt that the factor $(\log^+ \log^+ |f|)^{k-1}$ in (2) cannot be omitted. This factor is analogous to the logarithmic factor in the theorem on the strong differentiability of multiple integrals and the latter is known to be indispensable (see [2], [5]).

The purpose of this brief note is to consider the case of the hypersphere S. It turns out that the iterated logarithm does not enter there and the situation resembles very much the Nevanlinna—Ostrowski theorem.

Theorem. *Let $f(z_1, z_2, \dots, z_k)$ be regular in the hypersphere S and suppose that the integral*

$$\int_{\sigma_r} \log^+ |f(r_1 e^{i\theta_1}, \dots, r_k e^{i\theta_k})| \, d\sigma_r \tag{3}$$

is bounded, where σ_r denotes the boundary of the hypersphere $r_1^2 + r_2^2 + \dots + r_k^2 \leq r^2$ and $d\sigma_r$ is the element of volume of this boundary. Then for almost every point $(z_1^0, z_2^0, \dots, z_k^0)$ of σ_1 the function f has a finite non-tangential limit.

The proof is easy, if one uses a recent and important result of CALDERÓN (see [1]) concerning boundary values of functions harmonic in a hypersphere. This result may be stated as follows. Suppose that $u(\xi_1, \xi_2, \dots, \xi_n)$ is a real-valued harmonic function of the real variables $\xi_1, \xi_2, \dots, \xi_n$ in the hypersphere

$$\xi_1^2 + \xi_2^2 + \dots + \xi_n^2 < 1.$$

Suppose that for every point p_0 of a set E situated on the boundary of this hypersphere there is a (finite) cone with vertex at p_0, with axis along the radius terminating at p_0, and such that u is bounded in this cone. Then at almost every point of E the function u has a finite non-tangential limit. (For $n = 2$, this is an old result of PRIVALOFF (see [3]), who proved it by the method of conformal mapping. For $n > 2$, however, the proof requires a totally different idea).

Let us now assume the boundedness of the integral (3) for $r < 1$. The function $\log^+ |f(x_1 + iy_1, \dots, x_k + iy_k)|$ is in S a subharmonic function of the $2k$ variables $x_1, y_1, \dots, x_k, y_k$. The proof of this is essentially the same as in the familiar case $k = 1$. First of all, $\log^+ |f|$ is continuous in S. It is therefore enough to show that for every point p_0 in S, and for every sufficiently small sphere S' with center at p_0, the value of $\log^+ |f|$ at p_0 does not exceed the average of $\log^+ |f|$ taken on the boundary of S'. If f vanishes at p_0, this is immediate, since $\log^+ |f|$ is non-negative. If f is distinct from 0 at p_0, then in a sufficiently small neighborhood of p_0 the function $\log |f|$ is harmonic in each pair of the variables x_i, y_i, and so also is harmonic in all the variables $x_1, y_1, \dots, x_k, y_k$. Hence $\log^+ |f| = \mathrm{Max}\{0, \log |f|\}$ is subharmonic in that neighborhood, so that the required inequality is satisfied. Thus $\log^+ |f|$ is subharmonic in S.

Let $u_r(x_1, y_1, \dots, x_k, y_k)$ be the Poisson integral formed with the values of $\log^+ |f|$ on σ_r. Thus u_r is a non-negative harmonic function in the interior of σ_r and majorizes $\log^+ |f|$ there. By a familiar result from the theory of subharmonic functions, $u_r(x_1, \dots, y_k)$ is a non-decreasing function of r at every point $x_1, \dots, y_k$ (of course for the values of r such that $r^2 > x_1^2 + \dots + y_k^2$). By the theorem of HARNACK, u_r tends in S either to a harmonic function u, or to $+\infty$. The latter is impossible since

$$\frac{1}{\text{meas } \sigma_r}\int_{\sigma_r} \log^+|f|\,d\sigma_r \tag{4}$$

represents the value of u_r at the origin and is a bounded function of r [the boundedness of (4) follows from the fact that it is a non-decreasing function of r and from the boundedness of the integral (3)].

Thus the function $\log^+|f(z_1,\dots,z_k)|$ has a harmonic majorant $u(z_1,\dots,z_k)$ in S. This harmonic function being non-negative, it is the Poisson integral of a positive mass distributed over the boundary σ_1 of S. Hence (as in the case $k=1$) $u(z_1,\dots,z_k)$ has a non-tangential limit at almost every point of σ_1. Since $u \geqq \log^+|f|$, it follows that f is at any rate bounded as $(z_1,\dots,z_k)$ approaches non-tangentially almost every point of σ_1. Since the real and the imaginary part of f are real-valued harmonic functions, an application of CALDERÓN's result shows that f has a non-tangential limit at almost every point of σ_1. This completes the proof the theorem.

2. Let $f(z_1,\dots,z_k)$ be a function regular in S, and let p be a positive number. As in the case $k=1$, we say that f belongs to the class H^p, if the integral

$$\int_{\sigma_r} |f|^p\,d\sigma_r \tag{5}$$

remains bounded for $r<1$. Since $\log^+|f| \leqq |f|^p + \text{Const.}$, the boundedness of the integral (5) implies that of (3). Thus if f is of the class H^p, the non-tangential limit of f exists at almost every point of σ_1, and is of course of the class L^p over σ_1. We shall denote this limit also by f. Let $F(z_1,\dots,z_n)$ denote the upper bound of $|f|$ on the radius of S terminating at the point $(z_1,\dots,z_k) \in \sigma_1$. It has recently been proved (see [4]) that if $f \in H^p$, then $F \in L^p$ on σ_1.

From this we immediately deduce the following

Theorem. *If* $f(z_1, z_2,\dots,z_k) \in H^p$, *then*

$$\int_{\sigma_1} |f(z_1,\dots,z_k) - f(rz_1,\dots,rz_k)|^p\,d\sigma_1 \to 0 \quad \textit{as } r\to 1.$$

References.

1. A. P. CALDERÓN, On the behavior of harmonic functions at the boundary. To appear in the *Transactions of the American Math. Society.*
2. B. JESSEN, J. MARCINKIEWICZ and A. ZYGMUND, Note on the differentiability of multiple integrals, *Fundantenta Math.*, **25** (1935), pp. 217—234.
3. I. I. PRIVALOFF, *Cauchy's Integral* (Saratoff, 1919), in Russian.
4. H. E. RAUCH, Généralisation d'une proposition de Hardy et Littlewood et de théorèmes ergodiques qui s'y rattachent, *Comptes Rendus Acad. Sci. Paris,* **227** (1948), pp. 887—889.
5. S. SAKS, On the strong derivatives of functions of intervals, *Fundamenta Math.*, **25** (1935), pp. 235—252.
6. A. ZYGMUND, On the boundary values of functions of several complex variables, I. To appear in the *Fundamenta Math.*, **36**.

(Received August 17, 1949.)

Algebraic formulation of the problem of measure.

By M. H. STONE in Chicago.

Under a similar title, "Algebraische Fassung des Massproblems", ALFRED TARSKI has shown how the general problem of constructing an additive measure can be reduced to an interesting algebraic form[1]). In this paper we intend to carry TARSKI's reduction a little further[2]).

We consider, as TARSKI does, a non-void abstract set X with an arbitrary fixed triadic relation R, writing $R(x, y, z)$ to indicate that the elements x, y, z, of X are in the relation R. A non-negative real function μ defined on X will be called an R-additive measure if $R(x, y, z)$ implies $\mu(x) = \mu(y) + \mu(z)$. A suggestive notation consists in writing $x = y + z$ whenever $R(x, y, z)$; but the operation $+$ thus introduced need not be defined for all pairs (y, z), need not be single-valued, and need not enjoy any special algebraic properties. Illustrations and interpretations of these concepts may be found in TARSKI's paper[1]). A particularly interesting example arises in abstract geometry: let the geometrical structure of a space be defined in terms of a relation of congruence among its subsets, let X be a family of subsets of the given space, and let $R(x, y, z)$ if and only if x is the union of disjoint sets congruent respectively to y and to z; then an R-additive measure defined over X will in general have the characteristic properties of a geometrical content, being additive for the union of disjoint sets and invariant under the replacement of a set by a congruent one.

Since the real number system can be viewed as a rational-linear space, that is, as a linear space with the rational number system as coefficient-field, we may regard an R-additive measure as merely a special instance of an R-additive mapping of X into a rational-linear space. Our first step will be to analyze the structure of such general mappings. It is comparatively easy to show that any such mapping decomposes into a fixed mapping of X into

---

[1]) A. TARSKI, *Fundamenta Math.*, **31** (1938), pp. 47–66.

[2]) Our results were included in Colloquium Lectures delivered before the American Mathematical Society at Madison, Wisconsin, in September, 1939. They have also been presented in lectures at Harvard University (1945–46) and the University of Chicago (1946).

a certain fixed rational-linear space L_X and a rational-linear mapping of the space L_X. The problem of constructing an R-additive measure is thus reduced to the problem of constructing a rational-linear mapping of L_X into the real number system, subject to certain requirements of positivity. The latter problem can then be solved in terms of the theory of convex sets in a rational-linear space.

The rational-linear space L_X and the natural R-additive mapping T of X into L_X will now be constructed. Let L be the rational-linear space of all rational-valued functions f on X such that the set $(x; f(x) \neq 0)$ is finite. Let f_x designate that member of L defined by the relations $f_x(x) = 1$, $f_x(y) = 0$ when $y \neq x$. The mapping H: $x \to f_x$ carries X in a one-to-one manner onto a set of rationally-linearly independent elements in L. Let L_0 be the rational-linear subspace of L generated by the functions of the special form $f_x - f_y - f_z$ where $R(x, y, z)$; an element of L belongs to L_0 if and only if it is a rational-linear combination of a finite number of these special functions. Identification of the elements of L modulo L_0 in the standard way produces a rational-linear space $L_X = L - L_0$, and can be regarded as a rational-linear mapping of L on L_X. We shall denote this mapping as G and refer to it as the natural mapping of L on L_X. The mapping $T = GH$ is a mapping of X into L_X which is R-additive in the sense that $R(x, y, z)$ implies $Tx = Ty + Tz$: for $R(x, y, z)$ implies that the elements f_x and $f_y + f_z$ are to be identified modulo L_0 in as much as their difference is in L_0; and thus $Tx = G(Hx) = Gf_x = G(f_y + f_z) = Gf_y + Gf_z = G(Hy) + G(Hz) = Ty + Tz$.

We now have:

Theorem 1. *If M is a rational-linear space and A an R-additive mapping of X into M (in the sense that $R(x, y, z)$ implies $Ax = Ay + Az$), then there exists a rational-linear mapping S of L_X into M such that $A = ST$. Conversely, if S is any rational-linear mapping of L_X into M, then the mapping $A = ST$ of X into M is R additive.*

Proof. Consider the mapping AH^{-1}, which carries f_x into Ax. Since every element of L is a rational-linear combination of the rationally-linearly independent elements f_x, this mapping has a unique rational-linear extension U which maps L into M: U carries the element $\sum_{k=1}^{n} \alpha_k f_{x_k}$ of L into the element $\sum_{k=1}^{n} \alpha_k Ax_k$. Now $R(x, y, z)$ implies $U(f_x - f_y - f_z) = Uf_x - Uf_y - Uf_z = Ax - Ay - Az = 0$. It follows that U carries every element of the rational-linear subspace L_0 into the element 0 of M. Consequently U can be decomposed as a product $U = SG$ where S is a rational-linear mapping of $L_X = L - L_0$ into M and G is the natural mapping of L on L_X. We now have $A = UH = SGH = ST$, as we wished to prove. The converse statement is trivial.

Corollary. *If M is the real number system, considered as a rational-linear space, then the mapping A of Theorem 1 is an R-additive measure if and only if the mapping S is a rational-linear real-valued functional which assumes only non-negative values on the subset $T(X)$ of L_X.*

From Theorem 1 and its corollary, we see that the problem of finding the R-additive measures for X has been reduced to an algebraic problem about rational-linear spaces. We shall now discuss the latter problem in some detail. Let V be a rational-linear space with elements $u, v, w, \ldots$; and let V_0 be an arbitrary non-void subset of V. Our problem, to state it in one way, is to construct a real-valued function λ on V with the properties

$$(1)\quad \lambda(u+v)=\lambda(u)+\lambda(v), \qquad (2)\quad \lambda(u)\geqq 0 \text{ for } u\in V_0.$$

To eliminate trivial solutions we shall also require that

$$(3)\quad \lambda(v_0)>0 \text{ for some } v_0\in V_0.$$

From (1) we infer at once, in a familiar way, that λ is rational-linear, having the property that $\lambda(\alpha u)=\alpha\lambda(u)$ for all rational α. Let us consider what implications the existence of such a function λ may have for the relative positions of v_0 and V_0 in the rational-linear geometry of V. For this purpose we shall need some simple definitions. First, let us define a subset of V to be *convex* if, whenever it contains u and v, it also contains $\alpha u+\beta v$ where α and β are any *rational* numbers such that $\alpha\geqq 0$, $\beta\geqq 0$, $\alpha+\beta=1$. Similarly, let us define a subset of V to be a *cone with vertex* u_0 if, whenever it contains u, it also contains $\alpha u+\beta u_0$ where α and β are any *rational* numbers such that $\alpha\geqq 0$, $\alpha+\beta=1$. A convex cone with vertex at the origin is thus characterized by the property that, wheneve it contains u and v, it also contains $\alpha u+\beta v$ where α and β are any non-negative *rational* numbers. *In the sequel we shall use the term "cone" to mean always a convex cone with vertex at the origin,* no other type of cone being required for our purposes. Any non-void subset of V is contained in a smallest cone, consisting of all the linear combinations with non-negative rational coefficients of the element of the given subset. Finally we shall define a point u_0 beloging to a set U to be *internal* to U if for each $u\neq u_0$ the element $\alpha u+\beta u_0$, where α and β are rational numbers with $\alpha\geqq 0$, $\beta\geqq 0$, $\alpha+\beta=1$, belongs to U not merely for $\alpha=0$ but also for all sufficiently small $\alpha>0$. Geometrically this means that every straight line through u_0 has in common with U a certain segment on which u_0 lies but of which it is not an end-point. Consider now the smallest cone $C(V_0)$ containing V_0, and the set $K=(u;\lambda(u)<0)$. It is clear that λ is non-negative on $C(V_0)$ as well as on V_0, and hence that K is disjoint from $C(V_0)$. At the same time K is obviously a convex set containing $-v_0$. Indeed, it is easy to see that $-v_0$ is internal to K: for if $u\neq -v_0$ we have $\lambda(\alpha u-\beta v_0)=\alpha\lambda(u)-\beta\lambda(v_0)<0$ for all sufficiently small rational $\alpha>0$ when $\alpha+\beta=1$. These necessary properties can now be shown to be sufficient as well:

Theorem 2. *A necessary and sufficient condition for the existence of a function λ with properties* (1), (2), (3) *above is that* $-v_0$ *be internal to some convex set* K *disjoint from the smallest cone containing* V_0.

Proof. Only the sufficiency remains to be proved. ZORN's maximal principle[3]), is applicable to the system of cones containing V_0 and disjoint from K. Thus there is a maximal such cone, say S_0. An important property of S_0 is that it must contain at least one of the two elements u and $-u$, $u \in V$. To prove this, let us observe that the smallest cone containing the element u and the maximal cone S_0 consists of all elements of the form $\alpha u + v$ where α is a non-negative rational number and $v \in S_0$. Hence if u is not in S_0 this cone must have a point in common with K because of the maximality of S_0; in other words, there exist a positive rational α and an element v in S_0 such that $\alpha u + v \in K$. Similarly, if $-u$ is not in S_0 there exist a positive rational β and an element w in S_0 such that $-\beta u + w \in K$. If neither u nor $-u$ is in S_0 we obtain a contradiction as follows: putting $\alpha' = \beta/(\alpha+\beta)$, $\beta' = \alpha/(\alpha+\beta)$, we see that these positive rational numbers have 1 as sum; the convexity of K thus implies $\alpha' v + \beta' w = \alpha'(\alpha u + v) + \beta'(-\beta u + w) \in K$; but on the other hand, the convexity of S_0 implies $\alpha' v + \beta' w \in S_0$. In particular we see that the set $-K = (u\,;\, -u \in K)$ is contained in S_0 and obviously contains v_0 as an internal point. Hence S_0 also contains v_0 as an internal point. A second application of ZORN's principle provides us now with a maximal cone S containing S_0 and excluding $-v_0$. Clearly S, like its subset S_0, contains at least one of u and $-u$, $u \in V$. Since $U = (u\,;\, u \in S,\ -u \in S)$ is obviously a rational-linear subspace of V, we may effect the identification of elements in V modulo U so as to obtain the rational-linear space $W = V - U$. The natural mapping of V on W will be designated as λ. The space W can be ordered by putting $\lambda(u) < \lambda(v)$ if and only if $v - u \in S$, $u - v$ non $\in S$. In this ordering we have $\lambda(v_0) > 0$, $\lambda(u) \geqq 0$ for $u \in S$. Thus, if we can show that the order in W is archimedean, we can identify W with a rational linear subspace of the real number system, ordered in the standard way; and we can identify λ as the function which we desired to construct. Hence all that remains for us to do is to show that $\lambda(u) > 0$ implies the existence of positive rational numbers α and β for which $\alpha\lambda(v_0) \leqq \lambda(u) \leqq \beta\lambda(v_0)$. To find α, we observe that $-u$ non $\in S$ and hence that there exist a positive rational number γ and an element v in S for which $-\gamma u + v = -v_0$, in accordance with the maximality of the cone S. Taking $\alpha = 1/\gamma$, we have $u - \alpha v_0 = \dfrac{1}{\gamma} v \in S$ and hence $\lambda(u) \geqq \lambda(\alpha v_0) = \alpha\lambda(v_0)$. To find β, we note that for sufficiently small positive

[3]) This principle asserts that any non-void partially ordered system in which every chain is bounded has at least one maximal element. In our case we consider the system of cones to be partially ordered by inclusion.

rational α we have $\alpha(-u)+(1-\alpha)v_0 \in S$ because v_0 is internal to S. Then taking $\beta=(1-\alpha)/\alpha$ we have $\beta v_0 - u \in S$ and hence $\beta\lambda(v_0) \geqq \lambda(u)$. This completes the proof [4]).

Inspection of the proof just given shows that we can state

Theorem 3. *For the existence of a function λ with the properties* (1), (2), (3) *above it is sufficient that the smallest cone containing V_0 contain v_0 as an internal point but exclude $-v_0$.*

Using Theorem 3 it is easy to see that under certain circumstances the problem of the existence of an R-additive measure for X becomes a purely combinatorial one. We have

Theorem 4. *Let x_0 be such an element of X that Tx_0 is an internal point of the smallest cone containing TX in L_X. Then a necessary and sufficient condition for the existence of an R-additive measure μ for X with $\mu(x_0)>0$ is that for every finite subset Y of X containing x_0 there exist an R-additive measure μ_Y with $\mu_Y(x_0)>0$.*

Proof. The necessity of the condition is trivial since we can put $\mu_Y(x)=\mu(x)$ for all $x\in Y$ when the R-additive measure μ is known to exist. To prove the condition sufficient, we derive from it the result that $-Tx_0$ does not belong to the smallest cone containing TX; and we can then apply Theorem 3. If $-Tx_0$ belongs to the indicated cone, we may observe that this cone is the image under the natural mapping G of the smallest cone containing HX in L and hence that the latter cone must contain an element f which is to be identified modulo L_0 with $Hx_0=-f_{x_0}$. We note that the function f is non-negative and that there exist rational numbers α_k and elements x_k, y_k, z_k in X with $R(x_k, y_k, z_k)$ such that $f_{x_0}+f=g=\sum_{k=1}^{n}\alpha_k(f_{x_k}-f_{y_k}-f_{z_k})\in L_0$. Let Y be the set consisting of $x_0, x_1, y_1, z_1, \ldots, x_n, y_n, z_n$ and those elements x, finite in number, for which $f(x)\neq 0$. The rational-linear space of rational-valued functions defined on Y is isomorphic to the rational-linear subspace of L generated by $HY\subset HX$; and the rational-linear space L_Y associated with Y can be identified with the rational-linear subspace of L_X generated by $TY\subset TX$. The existence of the R-additive measure μ_Y with $\mu_Y(x_0)>0$ implies, by Theorem 1, the existence of a rational-linear real function λ_Y on L_Y such that λ_Y is non-negative on TY and $\lambda_Y(Tx_0)>0$. Observing now that $Gf_{x_0}=Tx_0$, Gf is in the smallest cone containing TY, and $Gf_{x_0}+Gf=Gg=0$, we arrive at the absurd result $0<\lambda_Y(Tx_0)+\lambda_Y(Gf)=\lambda_Y(0)=0$.

---

[4]) In this presentation, we have taken advantage of a short paper of J. Dieudonné, Sur le théorème de Hahn—Banach, *Revue Scientifique,* **79** (1941), pp. 642—643, where the author treats the Hahn—Banach theorem in terms of the theory of convexity in a real linear space.

Hence we see that $-Tx_0$ is outside the smallest cone containing TX, as we asserted above. This completes the proof.

By way of conclusion we may make two remarks. In the first place, G. MOSTOW[5]) has observed that in terms of a certain "natural" topology for rational-linear spaces Theorem 2 may be given the equivalent form: the required function λ exists if and only if $-v_0$ is not a point of the smallest closed cone containing V_0. The topology of MOSTOW is that in which a set is said to be open if and only if each of its points is internal to some convex part of the set. We shall not pursue this remark further here. Our second remark is the rather obvious one that the general theory developed in TARSKI's paper and in this one needs to be tested on specific examples. The discussion in TARSKI's paper shows that the construction of the space L_X conceals apparently difficult combinatorial problems met in determining whether or not certain elements in L are to be indentified modulo L_0; and this fact suggests the difficulties which can be anticipated in trying to ascertain the relative positions of Tx_0 and TX in L_X in any concrete case. Furthermore, the obsɩrvation that in some simple cases, familiar to everyone, the R-additive measure for X is essentially unique leads to an inquiry as to the conditions on TX which will guarantee uniqueness in general terms.

(Received July 26, 1949.)

[5]) Mr. MOSTOW was a member of a Harvard class to which I presented Theorem 2 in 1945–46. I have since made systematic use of his topology in developing the theory convexity relative to an arbitrary ordered field (unpublished).

Double points of paths of Brownian motion in n-space.

By A. DVORETZKY, P. ERDŐS and S. KAKUTANI in Urbana, Illinois.

§ 1. Introduction.

Let $(\Omega, \mathcal{E}, \Pr)$ be a probability space, i. e. $\Omega = \{\omega\}$ is a set of elements ω, $\mathcal{E} = \{E\}$ is a Borel field of subsets E of Ω called "events", and $\Pr(E)$ is a countably additive measure defined on $\mathcal{E}$ with the normalization $\Pr(\Omega) = 1$. $\Pr(E)$ is called the "probability" of the event E.

A *one-dimensional Brownian motion* [cf. **3, 5, 6, 7**] is a real-valued function $x(t, \omega)$ of the two variables t and ω, defined for all non-negative real numbers t, $0 \leqq t < \infty$, and for all $\omega \in \Omega$, with the following properties:

(B_1) $x(0, \omega) \equiv 0$,

(B_2) for any real numbers s, t with $0 \leqq s < t < \infty$, $x(t, \omega) - x(s, \omega)$ is $\mathcal{E}$-measurable in ω and has a Gaussian distribution with mean value 0 and variance $t - s$, i. e. [1])

(1) $$E_{x, s, t, \alpha, \beta} \equiv \{\omega \mid \alpha < x(t, \omega) - x(s, \omega) < \beta\} \in \mathcal{E},$$

and

(2) $$\Pr(E_{x, s, t, \alpha, \beta}) = \frac{1}{\sqrt{2\pi(t-s)}} \int_{\alpha}^{\beta} e^{-\frac{u^2}{2(t-s)}}\, du$$

for any real numbers α, β with $-\infty < \alpha < \beta < \infty$,

(B_3) for any real numbers s_k, t_k $(k = 1, \ldots, p)$ with $0 \leqq s_1 < t_1 \leqq s_2 < t_2 \leqq \leqq \ldots \leqq s_p < t_p < \infty$, the functions $x(t_k, \omega) - x(s_k, \omega)$, $k = 1, \ldots, p$, are independent in the sense of probability theory, i. e.

(3) $$\Pr(\bigcap_{k=1}^{p} E_{x, s_k, t_k, \alpha_k, \beta_k}) = \Pi_{k=1}^{p} \Pr(E_{x, s_k, t_k, \alpha_k, \beta_k})$$

for any real numbers α_k, β_k with $-\infty < \alpha_k < \beta_k < \infty$, $k = 1, \ldots, p$.

An *n-dimensional Brownian motion* is an n-system of mutually independent one-dimensional Brownian motions, i. e. an n-system $\{x^i(t, \omega) \mid i = 1, \ldots, n\}$ of one-dimensional Brownian motions $x^i(t, \omega)$, $i = 1, \ldots, n$, with the property that

(4) $$\Pr(\bigcap_{i=1}^{n} E^i) = \Pi_{i=1}^{n} \Pr(E^i),$$

where E^i is any subset of Ω determined by $x^i(t, \omega)$, i. e. a subset of Ω which belongs to the Borel subfield $\mathcal{E}^i$ of $\mathcal{E}$ which is generated by $\{E_{x, s, t, \alpha, \beta} \mid 0 \leqq s < t < \infty, \quad -\infty < \alpha < \beta < \infty\}$, $\quad i = 1, \ldots, n$.

---

[1]) $\{\omega \mid \ldots\}$ denotes the set of all ω having the properties ..., and similarly in other cases.

If we consider $\mathbf{x}(t, \omega) = \{x^i(t, \omega) \mid i = 1, \ldots, n\}$ as a point in an n-dimensional Euclidean space R^n, then, for each fixed ω, $\mathbf{x}(t, \omega)$ can be considered as an R^n-valued function of t defined for $0 \leqq t < \infty$.

It is easy to see that this definition of an n-dimensional Brownian motion is independent of the choice of the rectangular coordinate system, i. e. it is invariant vis-à-vis rotations of the coordinate system.

It is assumed (cf. Doob [1]) that the Borel field $\mathcal{E}$ is already extended by adding null sets in such a way that the subset C of Ω consisting of all ω for which $\mathbf{x}(t,\omega)$ is a continuous function of t for $0 \leqq t < \infty$ is $\mathcal{E}$-measurable and satisfies $\Pr(C) = 1$.

For any $\mathbf{y} = \{y^1, \ldots, y^n\} \in R^n$ and for any $\omega \in \Omega$, let us put

(5) $L_{a,b}^{(n)}(\mathbf{y}; \omega) = \{\mathbf{y} + \mathbf{x}(t, \omega) \mid a \leqq t \leqq b\}, \quad 0 \leqq a < b < \infty,$

(6) $L_{a,\infty}^{(n)}(\mathbf{y}; \omega) = \{\mathbf{y} + \mathbf{x}(t, \omega) \mid a \leqq t < \infty\}, \quad 0 \leqq a < \infty,$

(7) $L^{(n)}(\mathbf{y}; \omega) = L_{0,\infty}^{(n)}(\mathbf{y}; \omega),$

(8) $L_{a,b}^{(n)}(\omega) = L_{a,b}^{(n)}(\mathbf{0}; \omega), \quad L_{a,\infty}^{(n)}(\omega) = L_{a,\infty}^{(n)}(\mathbf{0}; \omega), \quad L^{(n)}(\omega) = L^{(n)}(\mathbf{0}; \omega),$

where $\mathbf{y} + \mathbf{x}(t, \omega) = \{y^i + x^i(t, \omega) \mid i = 1, \ldots, n\}$. $L_{a,b}^{(n)}(\mathbf{y}; \omega)$ is called the *(a, b)-path* of the n-dimensional Brownian motion starting from $\mathbf{y}$ and $L^{(n)}(\mathbf{y}; \omega)$ is called the *path* of the n-dimensional Brownian motion starting from $\mathbf{y}$.

For almost all ω (i. e. for all $\omega \in C$), $L_{a,b}^{(n)}(\mathbf{y}; \omega)$ is a continuous image of a closed interval $[a, b] = \{t \mid a \leqq t \leqq b\}$, and is hence a compact subset of R^n.

$\mathbf{x}_0 = \{x_0^1, \ldots, x_0^n\} \in R^n$ is called a double point of $L_{a,b}^{(n)}(\mathbf{y}; \omega)$ [resp. of $L_{a,\infty}^{(n)}(\mathbf{y}; \omega)$], if there exists a pair of real numbers s, t with $a \leqq s < t \leqq b$ [resp. $a \leqq s < t < \infty$] such that $\mathbf{x}_0 = \mathbf{y} + \mathbf{x}(s, \omega) = \mathbf{y} + \mathbf{x}(t, \omega)$ (i. e. $x_0^i = y^i + x^i(s, \omega) = y^i + x^i(t, \omega)$, $i = 1, \ldots, n$). It is clear that $\mathbf{x}_0$ is a double point of $L_{a,b}^{(n)}(\mathbf{y}; \omega)$ [resp. $L_{a,\infty}^{(n)}(\mathbf{y}; \omega)$] if and only if $\mathbf{x}_0 - \mathbf{y}$ is a double point of $L_{a,b}^{(n)}(\mathbf{0}; \omega) = L_{a,b}^{(n)}(\omega)$ [resp. $L_{a,\infty}^{(n)}(\mathbf{0}; \omega) = L_{a,\infty}^{(n)}(\omega)$].

It is known that (i) [Lévy 6] in R^2, almost all paths $L^{(2)}(\omega)$ of a 2 dimensional Brownian motion have double points and (ii) [3] in R^5, almost all paths $L^{(5)}(\omega)$ of a 5-dimensional Brownian motion have no double points. (ii) evidently implies that almost all paths in R^n with $n \geqq 5$ have no double points. Thus the problem of double points of paths of an n-dimensional Brownian motion is unsettled only for the cases $n = 3, 4$. These cases do not yield to the methods used in proving (i) and (ii); it is the purpose of this paper to dispose of these undecided cases by showing that (iii) in R^3, almost all paths $L^{(3)}(\omega)$ have double points, while (iv) in R^4, almost all paths $L^{(4)}(\omega)$ have no double points.

The proof of these results will be given in § 3 and § 4 respectively.

Our proof is based on the notion of capacity which plays an important role in the theory of harmonic functions in R^n. The definition of capacity and the statement of those of its fundamental properties which we need in the proofs of § 3 and § 4 will be found in § 2.

§ 2. Capacity.

Let F be a compact subset of $R^n (n \geqq 3)$. Let $\mathfrak{M}(F)$ be the family of all countably additive measures $m(B)$ defined for all Borel subsets B of F with $m(F)=1$. Let us put

$$\lambda^{(n)}(F) = \inf \iint \frac{m(d\mathbf{x})\, m(d\mathbf{y})}{|\mathbf{x}-\mathbf{y}|^{n-2}}, \tag{9}$$

where $|\mathbf{x}|$ denotes the distance of $\mathbf{x}$ from the origin $\mathbf{0}$ of R^n, so that $|\mathbf{x}-\mathbf{y}|$ is the distance of $\mathbf{x}$ and $\mathbf{y}$ in R^n; the double integral is extended over $F \times F$, and inf denotes the infimum for all measures $m \in \mathfrak{M}(F)$. $\lambda^{(n)}(F)=\infty$ if and only if the double integral is ∞ for all $m \in \mathfrak{M}(F)$. The n-dimensional capacity $C^{(n)}(F)$ of F is defined by

$$C^{(n)}(F) = \begin{cases} [\lambda^{(n)}(F)]^{-\frac{1}{n-2}} & \text{if} \quad \lambda^{(n)}(F) < \infty, \\ 0 & \text{if} \quad \lambda^{(n)}(F) = \infty. \end{cases} \tag{10}$$

The notion of capacity is important in the theory of harmonic functions in R^n, where under a harmonic function $f(\mathbf{x})$ defined in a domain D of R^n we understand a real-valued function $f(\mathbf{x})$ with continuous second partial derivatives which satisfies

$$\Delta f(\mathbf{x}) \equiv \sum_{i=1}^{n} \left(\frac{\partial}{\partial x^i}\right)^2 f(\mathbf{x}) \equiv 0 \tag{11}$$

in D.

In this paper we need the following properties of the capacity:

(C_1) [Frostman 2] Let $F = \{\mathbf{x}(t) | a \leqq t \leqq b\} \subset R^n$ be the continuous image of a closed interval $[a, b] = \{t \mid a \leqq t \leqq b\}$ of real numbers through the mapping $t \to \mathbf{x}(t)$. (This mapping need not be one-to-one.) Then the n-dimensional capacity of F is positive if

$$\int_a^b \int_a^b \frac{ds\,dt}{|\mathbf{x}(t)-\mathbf{x}(s)|^{n-2}} < \infty. \tag{12}$$

(C_2) [Pólya—Szegő 9] For any compact subset F of R^n, let us put

$$\lambda_p^{(n)}(F) = \inf \frac{2}{p(p-1)} \sum_{1 \leqq i < j \leqq p} \frac{1}{|\mathbf{x}_j - \mathbf{x}_i|^{n-2}}, \tag{13}$$

where inf denotes the infimum for all p-systems $\{\mathbf{x}_1, \dots, \mathbf{x}_p\} \subset F$. Then

$$\lim_{p \to \infty} \lambda_p^{(n)}(F) = \lambda^{(n)}(F). \tag{14}$$

(C_3) [9] The union of a finite number of compact subsets of R^n each of which has zero n-dimensional capacity has again zero n-dimensional capacity.

(C_4) [2] In order that a compact subset F of R^n have positive n-dimensional capacity, it is necessary and sufficient that there exist a function $g(\mathbf{y})$ harmonic, positive and smaller than 1 in R^n-F, and satisfying $g(\mathbf{y}) \to 0$ as $|\mathbf{y}| \to \infty$.

We need also the following result:

L e m m a 1. *Let F be a compact subset of R^n $(n \geqq 3)$. For any $\mathbf{y} \in R^n - F$ let* *us put* $\Omega(\mathbf{y};F)=\{\omega \mid L^{(n)}(\mathbf{y};\omega)\cap F\neq\theta\}$.[2] *Then* $\Omega(\mathbf{y};F)\in\mathcal{E}$ *and* $\Pr[\Omega(\mathbf{y};F)]=f(\mathbf{y};F)$ *is a harmonic function of* $\mathbf{y}$ *defined in* R^n-F. *Furthermore,* (i) $f(\mathbf{y};F)\equiv 0$ *in* R^n-F *if* $C^{(n)}(F)=0$; (ii) $0<f(\mathbf{y};F)<1$ *in* R^n-F; *and* $f(\mathbf{y};F)\to 0$ *as* $|\mathbf{y}|\to\infty$ *if* $C^{(n)}(F)>0$.

In the two-dimensional case the situation is rather different: (i) is still valid, but if the two-dimensional (logarithmic) capacity[3]) of F is positive then $f(\mathbf{y};F)\equiv 1$. This result can be found in [4] and the method of proof used there yields also our Lemma 1 for $n\geqq 3$. This is due to the property (C_4) of the capacity which holds only for $n\geqq 3$.

§ 3. The 3-dimensional case.

Lemma 2. *Let* $0\leqq a<b<\infty$. *Then, for almost all* ω, *the* (a,b)-*path* $L^{(3)}_{a,b}(\omega)$ *of a* 3-*dimensional Brownian motion has positive* 3-*dimensional capacity.*

Proof. Due to property (C_1) of the capacity, it suffices to show that

$$\int_a^b\int_a^b \frac{ds\,dt}{|\mathbf{x}(t,\omega)-\mathbf{x}(s,\omega)|}<\infty \tag{15}$$

for almost all ω, and hence it suffices to show that

$$I=\int_\Omega d\omega\int_a^b\int_a^b \frac{ds\,dt}{|\mathbf{x}(t,\omega)-\mathbf{x}(s,\omega)|}<\infty. \tag{16}$$

It is easy to see [by (B_2) and (B_3) of § 1] that

$$\begin{aligned}\int_\Omega \frac{d\omega}{|\mathbf{x}(t,\omega)-\mathbf{x}(s,\omega)|}&=\left(\frac{1}{\sqrt{2\pi|t-s|}}\right)^3\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\exp\left(-\frac{u^2+v^2+w^2}{2|t-s|}\right)}{\sqrt{u^2+v^2+w^2}}\,du\,dv\,dw=\\ &=\left(\frac{1}{\sqrt{2\pi|t-s|}}\right)^3\int_0^\infty\frac{\exp\left(-\frac{r^2}{2|t-s|}\right)}{r}\cdot 4\pi r^2dr=\left(\frac{1}{\sqrt{2\pi|t-s|}}\right)^3\cdot 4\pi|t-s|=\sqrt{\frac{2}{\pi|t-s|}}\end{aligned} \tag{17}$$

and consequently, by the Fubini theorem,

$$I=\int_a^b\int_a^b ds\,dt\int_\Omega\frac{d\omega}{|\mathbf{x}(t,\omega)-\mathbf{x}(s,\omega)|}=\sqrt{\frac{2}{\pi}}\int_a^b\int_a^b\frac{ds\,dt}{\sqrt{|t-s|}}<\infty. \tag{18}$$

We can now prove our first main result:

Theorem 1. *In a* 3-*dimensional Brownian motion, almost all paths* $L^{(3)}(\omega)$ *have infinitely many double points.*

Proof. Let $0\leqq a<b<c<\infty$. By Lemma 2, almost all (a,b)-paths $L^{(3)}_{a,b}(\omega)$ have a positive 3-dimensional capacity. By Lemma 1 and by the property (B_3) of Brownian motion, it is easy to see that $\Pr\{\omega \mid L^{(3)}_{a,b}(\omega)\cap L^{(3)}_{c,\infty}(\omega)\neq\theta\}>0$. From this it follows that there exists a real number d with $c<d<\infty$ such

---

[2]) θ denotes the empty set.

[3]) Cf. e. g. R. Nevanlinna [8].

that $\Pr\{\omega \mid L^{(3)}_{a,b}(\omega) \cap L^{(3)}_{c,d}(\omega) \neq \theta\} = \delta > 0$. Let us put $a_k = a + kd$, $b_k = b + kd$, $c_k = c + kd$, $d_k = (k+1)d$, $k = 1, 2, \ldots$. Then $\Pr\{\omega \mid L^{(3)}_{a_k, b_k}(\omega) \cap L^{(3)}_{c_k, d_k}(\omega) \neq \theta\} = \delta > 0$, $k = 1, 2, \ldots$, and consequently (since the independence property (B_3) enables us to reproduce the standard argument of the zero or one law) $\Pr\{\omega \mid L^{(3)}_{a_k, b_k}(\omega) \cap L^{(3)}_{c_k, d_k}(\omega) \neq \theta$ for infinitely many $k\} = 1$.

Remark. It is easily seen from the proof that for all $0 \leqq a < b < \infty$ and for almost all ω the (a, b) path $L^{(3)}_{a,b}(\omega)$ has infinitely many double points. Thus if we count only the doube points for which $0 < t - s < \delta$ where δ is an arbitrarily small positive number, then again almost all paths $L^{(3)}(\omega)$ have infinitely many such double points. Similarly, for any arbitrarily large $\Delta < \infty$, almost all paths $L^{(3)}(\omega)$ have infinitely many double points with $t - s > \Delta$. (Of course, the probability that $L^{(3)}_{a,b}(\omega)$ have such double points is always smaller than 1; it is zero if $\Delta \leqq b - a$ and positive otherwise.)

§ 4. The 4-dimensional case.

Lemma 3. *Let $0 \leqq a < b < \infty$. Then for almost all ω, the (a, b)-path $L^{(4)}_{a,b}(\omega)$ of a 4-dimensional Brownian motion has zero 4-dimensional capacity.*

Proof. By the uniform Lipschitz property of Brownian motion [Lévy **5**, § 52, pp. 166—173], there exist a finite constant M and a positive number $\delta(a, b, \omega)$ with $0 < \delta(a, b, \omega) < 1$ such that for almost all ω

(19) $$|\mathbf{x}(t, \omega) - \mathbf{x}(s, \omega)| < M \sqrt{|t - s| \log 1/|t - s|}$$

holds for all s and t with $a \leqq s < t \leqq b$ and $t - s < \delta(a, b, \omega)$. Since the closed interval $[a, b]$ is a union of a finite number of closed intervals of length $< \delta(a, b, \omega)$, the property (C_3) of the capacity inplies that it is sufficient to show that $L^{(4)}_{a,b}(\omega)$ has zero 4-dimensional capacity whenever $b - a \leqq 1$ and (19) is satisfied for all s, t with $a \leqq s < t \leqq b$. Thus, by property (C_2) of the capacity it suffices to prove

Lemma 4. *If we put*

(20) $$\lambda_p = \inf \frac{2}{p(p-1)} \sum_{1 \leqq i < j \leqq p} \frac{1}{|t_j - t_i| \log 1/|t_j - t_i|},$$

where inf denotes the infimum for all p-systems $\{t_1, \ldots, t_p\}$ of real numbers t_i $(i = 1, \ldots, p)$ such that $0 \leqq t_1 < \ldots < t_p < 1$, then

(21) $$\lim_{p \to \infty} \lambda_p = \infty.$$

Proof. Let N_m be the number of pairs (t_i, t_j) such that $2^{-m} \leqq t_j - t_i < 2^{-m+1}$, $m = 1, 2, \ldots$. Then

(22) $$N_m = \frac{1}{2} p(p-1)$$

and

(23) $$\frac{2}{p(p-1)} \sum_{1 \leqq i < j \leqq p} \frac{1}{|t_j - t_i| \log 1/|t_j - t_i|} \geqq \frac{2}{p(p-1)} \sum_{m=1}^{\infty} \frac{N_m}{2^{-m+1} \log 2^m} = \frac{1}{p(p-1) \log 2} \sum_{m=1}^{\infty} \frac{2^m N_m}{m}.$$

On the other hand, if we denote by $N_{m,k}$ the number of t satisfying $(k-1)2^{-m} \leqq t_i < k2^{-m}$, $k=1,\ldots,2^m$, then

$$\sum_{k=1}^{2^m} N_{m,k} = p \tag{24}$$

and

$$\sum_{l=m+1}^{\infty} N_l \geqq \sum_{k=1}^{2^m} \frac{1}{2} N_{m,k}(N_{m,k}-1). \tag{25}$$

This follows from the fact that $(k-1)2^{-m} \leqq t_i < t_j < k2^{-m}$ implies $t_j - t_i < 2^{-m}$. Consequently, by the Schwarz inequality,

$$\begin{aligned} N_m^* &\equiv \sum_{l=m+1}^{\infty} N_l \geqq \frac{1}{2}\left\{\sum_{k=1}^{2^m} N_{m,k}^2 - \sum_{k=1}^{2^m} N_{m,k}\right\} \geqq \\ &\geqq \frac{1}{2}\left\{\left(\sum_{k=1}^{2^m} N_{m,k}\right)^2 \Big/ 2^m - \sum_{k=1}^{2^m} N_{m,k}\right\} = \frac{1}{2}\left(\frac{p^2}{2^m} - p\right) \geqq \frac{p^2}{2^{m+2}}, \end{aligned} \tag{26}$$

where the last inequality holds for those m which satisfy $2^{m+1} \leqq p$, i.e. for $m \leqq m_p \equiv \left[\frac{\log p}{\log 2}\right] - 1$.

Consequently, by Abel's transformation, we have

$$\begin{aligned} \sum_{m=1}^{\infty} \frac{2^m N_m}{m} &= \sum_{m=1}^{\infty} \frac{2^m (N_{m-1}^* - N_m^*)}{m} = 2N_0^* + \sum_{m=1}^{\infty}\left(\frac{2^{m+1}}{m+1} - \frac{2^m}{m}\right) N_m^* \geqq \\ &\geqq \sum_{m=2}^{\infty} \frac{m-1}{m(m+1)} 2^m N_m^* \geqq \frac{1}{3}\sum_{m=2}^{\infty} \frac{2^m N_m^*}{m} \geqq \frac{1}{3}\sum_{m=2}^{m_p} \frac{2^m}{m}\frac{p^2}{2^{m+2}} = \\ &= \frac{p^2}{12}\sum_{m=2}^{m_p}\frac{1}{m} \geqq \frac{p^2}{12}[\log(m_p+1) - \log 2] \geqq \frac{p^2}{12}(\log\log p - 2\log 2) \end{aligned} \tag{27}$$

which, together with (20) and (23), imply

$$\lambda_p \geqq \frac{\log\log p}{12\log 2} - \frac{1}{6} \longrightarrow \infty \tag{28}$$

as $p \longrightarrow \infty$.

From this it is easy to deduce our last result:

Theorem 2. *In a 4-dimensional Brownian motion, almost all paths $L^{(4)}(\omega)$ have no double points.*

Proof. In view of (B_3), it suffices to show that, for any rational numbers a, b, c, d, with $0 \leqq a < b < c < d < \infty$, we have $\Pr\{\omega \mid L_{a,b}^{(4)}(\omega) \cap L_{c,d}^{(4)}(\omega) \neq \theta\} = 0$. But this last fact is an easy consequence of Lemmas 1 and 3.

Bibliography.

1. J. L. Doob, Stochastic processes depending on a continuous parameter, *Transactions American Math. Soc.*, **42** (1937), pp. 107—140.
2. O. Frostman, Potentiel d'équilibre et capacité des ensembles avec quelques applications à la théorie des fonctions, *Meddelanden Lunds Univ. Mat. Sem.*, **3** (1935).

3. S. Kakutani, On Brownian motion in n-space, *Proceedings Academy Tokyo,* **20** (1944), pp. 648–652.
4. S. Kakutani, Two-dimensional Brownian motion and harmonic functions, *ibidem,* **20** (1944), pp. 706–714.
5. P. Lévy, *Théorie de l'addition des variables aléatoires* (Paris, 1937).
6. P. Lévy, Le mouvement brownien plan, *American Journal of Math.,* **62** (1940), pp. 487–550.
7. P. Lévy, *Processus stochastiques et mouvements browniens* (Paris, 1948).
8. R. Nevanlinna, *Eindeutige analytische Funktionen* (Berlin, 1936).
9. G. Pólya—G. Szegő, Über den transfiniten Durchmesser (Kapazitätskonstante) von ebenen und räumlichen Punktmengen, *Journal für die reine und angewandte Math.,* **165** (1931), pp. 4—49.
10. N. Wiener, Generalized harmonic analysis, *Acta Math.,* **55** (1932), pp. 117—258.

(Received September 8, 1949.)

Une remarque sur les approximations diophantiennes linéaires.

Par VOJTĚCH JARNÍK à Praha.

Tous les nombres de cette Note sont réels. En particulier, les minuscules a, b, c désignent des nombres entiers, n, m, k, r des nombres entiers positifs. Soit $S=\{\theta_1, \ldots, \theta_r\}$ un système de r nombres réels. Je vais donner quelques résultats simples concernant la solution approximative de l'équation

$$(1) \qquad a_1\theta_1+a_2\theta_2+\ldots+a_r\theta_r+a_0=0, \quad |a_1|+|a_2|+\ldots+|a_r|>0$$

en nombres entiers. Pour caractériser le degré de précision avec laquelle cette équation peut être résolue, introduisons pour $t \geqq 1$ la fonction

$$(2) \qquad \psi_S(t)=\psi(t)=\min|a_1\theta_1+\ldots+a_r\theta_r+a_0|$$

où le minimum est pris pour a_i $(i=0, 1, 2, \ldots, r)$ variables sous la condition $0<\max(|a_1|,\ldots,|a_r|)\leqq t$. Si (1) n'a aucune solution, c'est-à-dire si l'on a $\psi(t)>0$ pour chaque $t\geqq 1$, nous allons dire que S est un système indépendant. On sait que $t^r\psi(t)<1$ pour chaque $t\geqq 1$. En particulier, pour $r=1$ on sait que si θ_1 est irrationnel (c'est-à-dire $\{\theta_1\}$ indépendant), on a toujours

$$0\leqq \liminf_{t\to+\infty} t\psi(t)\leqq\frac{1}{\sqrt{5}},\ ^{1)} \qquad \frac{1}{2}\leqq\limsup_{t\to+\infty} t\psi(t)\leqq 1\,;\ ^{2)}$$

donc, $\lim t\psi(t)$ n'existe pas. Pour $r>1$, on a un résultat analogue: une limite *positive* de $t^r\psi(t)$ ne peut pas exister. Plus précisément:

Théorème 1. *Si* $r\geqq 1$, $S=\{\theta_1,\ldots,\theta_r\}$,

$$\limsup_{t\to+\infty} t^r\psi_S(t)=B, \quad \liminf_{t\to+\infty} t^r\psi_S(t)=A>0,$$

on a $B>A$*; d'une manière plus précise,*

$$(3) \qquad \left(\frac{B}{A}\right)^{2^{r+1}}\geqq 2.$$

---

[1] A. HURWITZ. Voir p. ex. J. F. KOKSMA, *Diophantische Approximationen* (Berlin, 1936), p. 31, Satz 13 a.

[2] A. CHINČIN (KHINTCHINE). Voir p. ex. J. F. KOKSMA, l. c. [1], p. 36, Satz 24.

Mais, contrairement au cas $r=1$, on peut avoir $\lim t^r \psi_S(t)=0$, si $r>1$.[3]) Pour classifier, dans ce cas, l'ordre maximum et minimum de $\psi_S(t)$, définissons:

Soit $\alpha(S)=\alpha$ la borne supérieure de tous les γ pour lesquels

$$\limsup_{t\to+\infty} t^\gamma \psi_S(t) < +\infty;$$

soit $\beta(S)=\beta$ la borne supérieure de tous les γ pour lesquels

$$\liminf_{t\to+\infty} t^\gamma \psi_S(t) < +\infty.$$

On a donc $r \leqq \alpha \leqq \beta \leqq +\infty$. En outre on a, pour $r=2$, le théorème suivant:

Théorème 2. *Soit $S=\{\theta_1, \theta_2\}$ un système tel que $2<\alpha<+\infty$; alors $\beta \geqq \alpha(\alpha-1)>\alpha$.*

Donc, le cas $\alpha=\beta$ ne peut se présenter que si l'on a ou bien $\alpha=\beta=2$ ou bien $\alpha=\beta=+\infty$.

Remarque. Pour $r>2$ on a un théorème analogue, mais moins satisfaisant: Soit $S=\{\theta_1, \ldots, \theta_r\}$ $(r \geqq 2)$ un système tel que $(5r^2)^{r-1}<\alpha<+\infty$; alors $\beta \geqq \alpha^{\frac{r}{r-1}}-3\alpha>\alpha$. Et la borne donnée par cette inégalité est assez précise. En effet: Soit $r>1$. Alors il existe une suite de systèmes $S_n\{\theta_{1,n}, \ldots, \theta_{r,n}\}$ $(n=1, 2, \ldots)$ telle que l'on ait (en posant $\alpha_n=\alpha(S_n)$, $\beta_n=\beta(S_n)$)

$$\beta_n=n+O\left(n^{\frac{r-1}{r}}\right),\quad \beta_n=\alpha_n^{\frac{r}{r-1}}+O(\alpha_n) \text{ pour } n\to+\infty.$$

Les démonstrations de ces résultats seront publiées ailleurs. Théorème 2 est une conséquence immédiate du

Théorème 3. *Soit $S=\{\theta_1, \theta_2\}$ un système indépendant. Soit $\varphi(t)$ une fonction continue et décroissante pour $t \geqq 1$, $\lim_{t\to+\infty} t\varphi(t)=0$. Alors: Si l'on a $\psi_S(t)<\varphi(t)$ pour chaque $t>\tau_0$, il existe à chaque T un $t>T$ tel que*

$$\psi_S(t)<\varphi\left(\frac{1}{6t\,\varphi(t)}\right).$$

En posant $\varphi(t)=t^{-\gamma}$, $1<\gamma<\alpha$, on obtient aussitôt le Théorème 2. D'autre part, en posant $\varphi(t)=\lambda t^{-2}$ $(\lambda>B)$, on obtient

Théorème 4. *Si $r=2$, on peut remplacer, dans le Théorème 1, l'inégalité* (3) *par l'inégalité $A \leqq 36B^3$ (qui est plus précise que* (3) *si B est petit).*

Passons maintenant aux démonstrations très simples des Théorèmes 1, 3.

Soit $S=\{\theta_1, \ldots, \theta_r\}$ un système indépendant. La fonction $\psi(t)$ est donc positive et non croissante pour $t \geqq 1$, constante dans chaque intervalle

[3]) On a même le résultat suivant: Soit $r>1$; soit $\varphi(t)$ une fonction croissante et positive pour $t \geqq 1$. Alors il existe un système indépendant $S=\{\theta_1, \ldots, \theta_r\}$ tel que $\lim \varphi(t)\,\psi_S(t)=0$ (A. Chinčin). Voir p. ex. J. F. Koksma, l. c. [1]), p. 69, Satz 8.

$m \leqq t < m+1$ et l'on a $\lim\limits_{t\to+\infty} \psi(t) = 0$. Il existe donc une suite infinie de nombres naturels

(4) $$1 = t_1 < t_2 < t_3 < \ldots$$

telle que $\psi(t) = \psi(t_n)$ pour $t_n \leqq t < t_{n+1}$, $\psi(t_{n+1}) < \psi(t_n)$. À chaque n, il existe $r+1$ nombres $a_{i,n}$ $(i = 0, 1, \ldots, r)$ tels que

(5) $$a_{1n}\theta_1 + \ldots + a_{rn}\theta_r + a_{0n} = \psi(t_n), \quad \max(|a_{1n}|, \ldots, |a_{rn}|) = t_n.$$

Evidemment, le plus grand diviseur commun est égal à un:

(6) $$(a_{1n}, a_{2n}, \ldots, a_{rn}, a_{0n}) = 1.$$

Démonstration du théorème 1. Soit $0 < A' < A \leqq B < B' < +\infty$; donc

(7) $$A' t^{-r} < \psi(t) < B' t^{-r},$$

si t est assez grand. Pour $t_n \leqq t < t_{n+1}$ on a $\psi(t) = \psi(t_n)$; donc, pour $t = t_n$ d'une part et pour $t \to t_{n+1}$ d'autre part on obtient de (7), si n est assez grand,

(8) $$A' t_n^{-r} < \psi(t_n) \leqq B' t_{n+1}^{-r}.$$

Posons $k = 2^{r+1}$ et considérons les $k+1$ nombres (différents deux-à-deux)

$$\psi(t_{n+i}) = a_{1,n+i}\theta_1 + \ldots + a_{r,n+i}\theta_r + a_{0,n+i} \qquad (i = 0, 1, \ldots, k).$$

Ces nombres sont situés dans l'intervalle $0 < \xi \leqq \psi(t_n)$. Parmi ces nombres il y en a donc deux dont la différence $\varDelta$ est positive et plus petite que $\psi(t_n)/k$. On a $\varDelta = c_1\theta_1 + \ldots + c_r\theta_r + c_0$; puisque $0 < \varDelta < 1$, on a

$$0 < \max(|c_1|, \ldots, |c_r|) < 2t_{n+k},$$

donc $\psi(2t_{n+k}) \leqq \varDelta < k^{-1}\psi(t_n)$. Mais (8) donne

$$\psi(2t_{n+k}) \geqq A' 2^{-r} t_{n+k}^{-r} > 2^{-r} A' \left(\frac{A'}{B'}\right)^{k-1} t_{n+1}^{-r}, \quad k^{-1}\psi(t_n) \leqq k^{-1} B' t_{n+1}^{-r},$$

donc

$$\left(\frac{A'}{B'}\right)^k < \frac{2^r}{k} = \frac{1}{2}.$$

Pour $A' \to A$, $B' \to B$ on obtient le théorème.

Démonstration du théorème 3. Soit $r = 2$; écrivons $\theta, \eta, a_n, b_n, c_n$ au lieu de $\theta_1, \theta_2, a_{1n}, a_{2n}, a_{0n}$. On a

(9) $$\max(|a_n|, |b_n|) = t_n, \quad a_n\theta + b_n\eta + c_n = \psi(t_n),$$

(10) $$(a_n, b_n, c_n) = 1.$$

Supposons que pour tous les t assez grands on ait

(11) $$\varphi\left(\frac{1}{6t\varphi(t)}\right) \leqq \psi(t) < \varphi(t);$$

il faut en déduire une contradiction. Pour $t_n \leqq t < t_{n+1}$ on a $\psi(t) = \psi(t_n)$. Pour $t = t_n$ d'une part et pour $t \to t_{n+1}$ d'autre part on obtient donc de (11)

(12) $$\varphi\left(\frac{1}{6t_n\varphi(t_n)}\right) \leqq \psi(t_n) \leqq \varphi(t_{n+1}),$$

si n est assez grand. Il s'ensuit que

(13) $$t_{n+1} \leqq \frac{1}{6t_n\varphi(t_n)}.$$

En effet, dans le cas contraire, on aurait

$$\psi(t_n) \leqq \varphi(t_{n+1}) < \varphi\left(\frac{1}{6t_n\varphi(t_n)}\right)$$

ce qui contredit à (12).

Posons

$$D_n = \begin{vmatrix} a_n & b_n & c_n \\ a_{n+1} & b_{n+1} & c_{n+1} \\ a_{n+2} & b_{n+2} & c_{n+2} \end{vmatrix}, \quad E_n = \begin{vmatrix} a_n & b_n \\ a_{n+1} & b_{n+1} \end{vmatrix}.$$

J'affirme que $D_n = 0$, $E_n \neq 0$ si n est assez grand. En effet: En multipliant les équations

(14) $$a_{n+i}\theta + b_{n+i}\eta + c_{n+i} = \psi(t_{n+i}) \qquad (i = 0, 1, 2)$$

resp. par $a_{n+1}b_{n+2} - a_{n+2}b_{n+1}$, $a_{n+2}b_n - a_n b_{n+2}$, $a_n b_{n+1} - a_{n+1}b_n$ et en ajoutant, on obtient

$$|D_n| < 6t_{n+1}t_{n+2}\psi(t_n) \leqq 6t_{n+1}t_{n+2}\varphi(t_{n+1}),$$

donc, d'après (13), $|D_n| < 1$, $D_n = 0$.

D'autre part, si $E_n = 0$, on déduirait de (14) pour $i = 0, 1$:

$$|a_n c_{n+1} - a_{n+1}c_n| = |a_n\psi(t_{n+1}) - a_{n+1}\psi(t_n)| < 2t_{n+1}\psi(t_n) \leqq 2t_{n+1}\varphi(t_{n+1}) < 1$$

si n est assez grand, donc $a_n c_{n+1} - a_{n+1}c_n = 0$ et de même $b_n c_{n+1} - b_{n+1}c_n = 0$, donc $a_{n+1} = \tau a_n$, $b_{n+1} = \tau b_n$, $c_{n+1} = \tau c_n$, $|\tau| > 1$, ce qui contredit à (10).

Choisissons un k qui va rester fixe de sorte que

(15) $$D_n = 0, \ E_n \neq 0 \text{ pour chaque } n \geqq k.$$

Considérons la matrice infinie

(16) $$\begin{matrix} a_k & b_k & c_k \\ a_{k+1} & b_{k+1} & c_{k+1} \\ a_{k+2} & b_{k+2} & c_{k+2} \\ \cdots & \cdots & \cdots \end{matrix}$$

En considérant trois lignes consécutives quelconques de cette matrice, on voit que la troisième ligne est une combinaison linéaire des deux premières lignes qui, à leur tour, sont linéairement indépendantes. Il s'ensuit qu'une ligne quelconque de la matrice (16) est une combinaison linéaire de ses deux premières lignes. Donc

(17) $$\begin{vmatrix} a_k & b_k & c_k \\ a_{k+1} & b_{k+1} & c_{k+1} \\ a_n & b_n & c_n \end{vmatrix} = 0 \text{ pour tout } n \geqq k.$$

En posant

$$A = a_k b_{k+1} - a_{k+1} b_k, \quad A_n = a_{k+1} b_n - a_n b_{k+1}, \quad B_n = a_n b_k - a_k b_n$$

et en désignant par $C_1, C_2, \ldots$ des nombres qui ne dépendent pas de n, on déduit des équations

$$a_i \theta + b_i \eta + c_i = \psi(t_i) \qquad (i = k, k+1, n)$$

et de (17) aussitôt

$$A\psi(t_n) + A_n \psi(t_k) + B_n \psi(t_{k+1}) = 0. \tag{18}$$

Ici $A = E_k$ ne dépend pas de n et l'on a

$$A \neq 0, \quad |A_n| < C_1 t_n, \quad |B_n| < C_1 t_n.$$

Si l'on avait $B_n = 0$, on aurait d'après (18) $A_n \neq 0$ (car $A \neq 0$), donc

$$|A\psi(t_n)| = |A_n \psi(t_k)| \geqq \psi(t_k),$$

ce qui est impossible si n est assez grand.

L'équation (18) et l'équation analogue avec $n+1$ au lieu de n donnent, si n est assez grand,

$$|B_{n+1} A_n - B_n A_{n+1}| = |A| \,.\, |B_{n+1} \psi(t_n) - B_n \psi(t_{n+1})| / \psi(t_k) <$$
$$< C_2 t_{n+1} \psi(t_n) \leqq C_2 t_{n+1} \varphi(t_{n+1}) < 1$$

(voir (12)), donc $B_{n+1} A_n - B_n A_{n+1} = 0$, $B_{n+1} \psi(t_n) - B_n \psi(t_{n+1}) = 0$, c'est-à-dire

$$(B_{n+1} a_n - B_n a_{n+1})\, \theta + (B_{n+1} b_n - B_n b_{n+1})\, \eta + (B_{n+1} c_n - B_n c_{n+1}) = 0,$$

où $B_n \neq 0$, $B_{n+1} \neq 0$. Mais θ, η étant un système indépendant, il s'ensuit

$$B_{n+1} a_n - B_n a_{n+1} = B_{n+1} b_n - B_n b_{n+1} = B_{n+1} c_n - B_n c_{n+1} = 0,$$

d'où $a_n b_{n+1} - a_{n+1} b_n = 0$, ce qui contredit à (15).

(Reçu le 17 septembre 1949.)

Uniform Distribution and Lebesgue Integration.

By J. F. KOKSMA in Amsterdam and R. SALEM in Cambridge, Mass.

1. If $u_1, u_2, \dots$ denotes a sequence of real numbers uniformly distributed modulo 1 and if $f(x)$ is a bounded Riemann-integrable function of the real variable x, with period 1, then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(u_n) = \int_0^1 f(t)\,dt.$$

It is obvious that the theorem becomes false if, instead of supposing that f is Riemann-integrable, we assume only that f is Lebesgue-integrable, since we can change arbitrarily the values of f at all points u_n (mod 1) without changing the integral.

A natural question to ask is whether for $f \in L$, the relation

$$(1) \qquad \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(x+u_n) = \int_0^1 f(t)\,dt$$

holds almost everywhere in x. If $u_n = \theta n$, where θ is any fixed irrational number, the relation (1) holds for almost all x, under the only assumption that $f \in L$. This result, due to KHINTCHINE[1]) is actually an instance of BIRKHOFF's ergodic theorem[2]), and one cannot expect a generalization of the argument to general uniformly distributed sequences.

Here, using an argument based on different ideas, we shall give some results of the type (1), confining ourselves to the case $f \in L^2$ and to certain types of sequences $\{u_n\}$.

If, instead of a result of the type (1) we consider convergence in mean, we can state the following general theorem[3]):

---

1) A. KHINTCHINE, Eine arithmetische Eigenschaft der summierbaren Funktionen, *Recueil Math. Moscou*, 41 (1934), pp. 11—13.

2) For litterature see 1).

3) This theorem, the proof of which is very simple, may be known but we did not find it in the litterature.

Theorem I. *Let $f(x) \in L^2$ be a function with period* 1 *and mean value zero, i. e.* $\int_0^1 f(x)\,dx = 0$. *Then, for any sequence* $\{u_n\}$ *uniformly distributed modulo* 1, *one has*

$$\lim_{N\to\infty} \int_0^1 \frac{1}{N} \left| \sum_{n=1}^{N} f(x+u_n) \right|^2 dx = 0.$$

Proof. Let $\sum_{-\infty}^{\infty} c_k e^{2\pi i k x}$ be the Fourier series of $f(x)$, $(c_0 = 0,\ c_{-k} = \bar{c}_k)$. Let us write

$$S_k = \frac{1}{N}\left(e^{2\pi i k u_1} + \ldots + e^{2\pi i k u_N}\right),$$

so that the integral considered in the theorem is equal to

$$2 \sum_{1}^{\infty} |c_k|^2 |S_k|^2$$

and, since $|S_k| \leqq 1$, does not exceed

$$2 \sum_{1}^{h} |c_k|^2 |S_k|^2 + 2 \sum_{h+1}^{\infty} |c_k|^2.$$

If we now choose h such that $\sum_{h+1}^{\infty} |c_k|^2 < \varepsilon$ $(\varepsilon > 0)$, and then N_0 such that $|S_k|^2 < \varepsilon$ for $k = 1, 2, \ldots, h$ and $N \geqq N_0$ the integral will not exceed

$$\varepsilon \left[\int_0^1 f^2\,dx + 2 \right]$$

for $N \geqq N_0$, which proves the theorem.

2. We are unable to state a result of the type (1) without making certain additional hypotheses on the function f and on the sequence $\{u_n\}$. (That some additional hypotheses, at least on the function f, are necessary, will be shown at the end of the paper, with the use of an argument due to ERDŐS).

Let again $f \in L^2$ have period 1 and mean value zero, so that

$$f(x) \sim \sum_{-\infty}^{\infty} c_k e^{2\pi i k x} \qquad (c_0 = 0,\ c_{-k} = \bar{c}_k).$$

Let us denote by $R(h)$ the remainder $\sum_{h+1}^{\infty} |c_k|^2$.

Let us now denote by $S(M, N, k)$ the sum

$$\sum_{n=M+1}^{M+N} e^{2\pi i k u_n} \qquad (M, N \text{ and } k \text{ being integers}).$$

We can state the following theorem:

Theorem II. *Let $f \in L^2$ have period 1 and mean value zero, and be such that $R(h) = O\left(\frac{1}{(\log h)^\alpha}\right)$ where $\alpha > 1$. Let $\{u_n\}$ be a sequence uniformly distributed modulo 1 such that*

$$|S(M, N, k)| \leqq \varDelta k^\varrho N^\sigma (M+N)^\tau \qquad (k \geqq 1,\ M \geqq 1,\ N \geqq 1),$$

where $\varDelta$, ϱ, σ, τ are constants such that $\sigma + \tau < 1$ and $\tau < 1/2$. Then, almost everywhere in x,

$$\lim_{N \to \infty} \frac{1}{N} [f(x+u_1) + \ldots + f(x+u_N)] = 0.$$

Remark. As $f(x)$ needs not be bounded, theorem II is applicable to certain periodical functions which are only improperly integrable in the sense of RIEMANN.

The proof depends on the following lemma, which is a particular case of a result of GÁL and KOKSMA[4]). We give here a proof somewhat different from the original one.

Lemma: *Let $\{f_\nu(x)\}$, $\nu = 1, 2, \ldots$ be a sequence of functions all belonging to L^p $(p > 1)$ in the interval $(0, 1)$. Let $\eta(N)$ be positive monotonic decreasing such that $\sum \frac{\eta(N)}{N} < \infty$. Suppose that for all $M \geqq 0$, $N \geqq 1$*

$$\int_0^1 \left| \sum_{\nu=M+1}^{M+N} f_\nu \right|^p dx \leqq C (M+N)^{p-\lambda} N^\lambda \eta(N)$$

where $\lambda > 1$. Then, for almost all x,

$$\lim_{N \to \infty} \frac{1}{N} (f_1 + f_2 + \ldots + f_N) = 0.$$

Proof of the lemma. Let n be a positive integer. By $\varDelta_k^{(h)}$ $(h = 1, 2, \ldots, 2^k)$ we denote any of the intervals (open on the left, closed on the right) obtained by the subdivision of the interval $(0, 2^n)$ in 2^k equal parts. By $S_k^{(h)}$ we denote the sum Σf_ν where ν takes all integral values contained in $\varDelta_k^{(h)}$.

Denoting by j any fixed integer such that $1 \leqq j \leqq 2^n$, and writing j in the dyadic system, we find that the interval $(0, j)$ is the sum of certain intervals $\varDelta_k^{(h)}$ where k takes at most once each value $0, 1, 2, \ldots, n$, and each h depends on the corresponding k. According to this

$$\sum_{\nu=1}^{j} f_\nu = \varepsilon_0 S_0^{(h_0)} + \ldots + \varepsilon_n S_n^{(h_n)},$$

where $\varepsilon_i = 0$ or 1.

[4]) I. S. GÁL and J. F. KOKSMA, Sur l'ordre de grandeur des fonctions sommables, *Comptes Rendus Acad. Sci. Paris*, **227** (1949), pp. 1321—1323. The complete proof of the general theorem will appear elsewhere.

Let θ be a positive number larger than 1, to be fixed later on, one has, using HÖLDER's inequality,

$$\left|\sum_{\nu=1}^{j} f_\nu\right|^p \leqq \left(\sum_{k=0}^{n} \frac{1}{\theta^k}\,\theta^k\,|S_k^{(h_k)}|\right)^p \leqq \left(\sum_{k=0}^{n} \frac{1}{\theta^{p'k}}\right)^{p-1} \left(\sum_{k=0}^{n} \theta^{pk}|S_k^{(h_k)}|^p\right),$$

where $\dfrac{1}{p}+\dfrac{1}{p'}=1$. Hence, for *all* j $(1 \leqq j \leqq 2^n)$ and *all* x

$$\left|\sum_{\nu=1}^{j} f_\nu\right|^p \leqq B \sum_h \sum_k \theta^{pk}|S_k^{(h)}|^p,$$

where $B=\left(\sum_{k=0}^{\infty} \dfrac{1}{\theta^{p'k}}\right)^{p-1}$, and the double summation is extended to $k=0$, $1, 2, \ldots, n$, and for each k to *all* values of h $(h=1, 2, \ldots, 2^k)$. Now, by hypothesis,

$$\int_0^1 |S_k^{(h)}|^p\,dx \leqq C 2^{n(p-\lambda)} 2^{(n-k)\lambda}\,\eta\,(2^{n-k}).$$

Hence

$$\int_0^1 \left|\sum_{\nu=1}^{j(x)} f_\nu\right|^p dx \leqq B \sum_{k=0}^{n} \theta^{pk}.\,2^k C 2^{n(p-\lambda)} 2^{(n-k)\lambda}\eta(2^{n-k}),$$

where we can suppose that the integer $j(x)$ is any measurable function of x. Supposing now $2^{n-1} < j(x) \leqq 2^n$, one has

$$\int_0^1 \left|\frac{\sum_{\nu=1}^{j(x)} f_\nu}{j(x)}\right|^p dx = O\left\{\frac{1}{2^{pn}} \sum_{k=0}^{n} \theta^{pk} 2^k 2^{n(p-\lambda)} 2^{(n-k)\lambda}\,\eta(2^{n-k})\right\}$$

$$= O\left\{\sum_{k=0}^{n} \frac{\theta^{pk}}{2^{(\lambda-1)k}}\,\eta(2^{n-k})\right\}.$$

Now fix θ such that $1 < \theta < 2^{\frac{\lambda-1}{p}}$ (which is possible since $\lambda > 1$), and put

$$\alpha = \frac{\theta^p}{2^{\lambda-1}} < 1\,;$$

one has

$$\sum_{k=0}^{n} \alpha^k \eta\,(2^{n-k}) = \sum_{0}^{\left[\frac{n}{2}\right]} + \sum_{\left[\frac{n}{2}\right]}^{n} = O\,(\eta(2^{n/2})) + O\,(\alpha^{n/2}),$$

and, remarking that the condition $\sum \dfrac{\eta(N)}{N} < \infty$ implies $\sum \eta\,(2^{n/2}) < \infty$, one has, writing

$$I_n = \int_0^1 \left|\frac{\sum_{\nu=1}^{j(x)} f_\nu}{j(x)}\right|^p dx \qquad (2^{n-1} < j(x) \leqq 2^n),$$

that $\Sigma I_n < \infty$. In other words,

$$\sum_{n=1}^{\infty} \int_0^1 \max_{2^{n-1} < j \leq 2^n} \left| \frac{\sum_{\nu=1}^{j} f_\nu}{j} \right|^p dx < \infty,$$

which implies

$$\sum_1^N f_\nu = o(N)$$

for almost all x.

Proof of Theorem II. Writing

$$T_{M,N} = \int_0^1 \left| \sum_{n=M+1}^{M+N} f(x+u_n) \right|^2 dx$$

one has, using the hypotheses of the theorem:

$$T_{M,N} = 2 \sum_{k=1}^{\infty} |c_k|^2 |S(M, N, k)|^2 \leq$$

$$\leq 2A^2 \sum_{k=1}^{h} |c_k|^2 k^{2\varrho} N^{2\sigma}(M+N)^{2\tau} + 2N^2 \sum_{k=h+1}^{\infty} |c_k|^2 \leq$$

$$\leq A' \left[h^{2\varrho} N^{2\sigma}(M+N)^{2\tau} + \frac{N^2}{(\log h)^\alpha} \right],$$

A' being a constant. Fix now an ε, positive, such that

$$2\varrho\varepsilon + 2\sigma + 2\tau < 2 \tag{2}$$

as is clearly possible since $\sigma + \tau < 1$, and take for h the integral part of N^ε. Then

$$T_{M,N} \leq C \left[N^{2\varrho\varepsilon+2\sigma}(M+N)^{2\tau} + \frac{N^2}{(\log N)^\alpha} \right],$$

C being a constant. Writing

$$T_{M,N} \leq C \left[\frac{(M+N)^{2\tau} N^{2-2\tau}}{N^{2-2\tau-2\varrho\varepsilon-2\sigma}} + \frac{N^2}{(\log N)^\alpha} \right],$$

one has by (2)

$$T_{M,N} \leq D \frac{(M+N)^{2\tau} N^{2-2\tau}}{(\log N)^\alpha},$$

D being a constant. Since $\tau < 1/2$, $\alpha > 1$, an application of the lemma (with $p = 2$) gives

$$\lim_{N \to \infty} \frac{1}{N} [f(x+u_1) + f(x+u_2) + \ldots + f(x+u_N)] = 0$$

for almost all x.

3. Applications. We propose now to give examples of sequences $\{u_n\}$ uniformly distributed (mod 1) for which the relation

$$|S(M, N, k)| \leqq A k^{\varrho} N^{\sigma} (M+N)^{\tau} \quad (\sigma+\tau<1,\ \tau<1/2)$$

is satisfied.

First Example. Let θ denote an irrational number of the type I, that is to say that for some constant $\eta > 2$, the inequality

$$\left|\theta - \frac{p}{q}\right| < \frac{1}{q^{\eta}}$$

has only a finite number of solutions in integers p and q, $(q>1)$. We can take, for instance, for θ any algebraic number; or any irrational number with bounded partial quotients. By a well known theorem the numbers which are not of the type I form a null set (BOREL).

Let now, r being an integer $\geqq 2$,

$$u_n = \theta n^r + \alpha_1 n^{r-1} + \ldots + \alpha_r,$$

where $\alpha_1, \ldots, \alpha_r$ are arbitrary real constants. We shall prove that for the sequence $\{u_n\}$ using the notations of theorem II, one has

$$(3) \qquad |S(M, N, k)| \leqq A k^{\varrho} N^{\sigma} \qquad (\sigma < 1)$$

so that theorem II is applicable to such a sequence.

In fact, this can be deduced from theorems of WEYL, VINOGRADOFF and others. As we do not need the modern results in their sharpest form, we make use, instead, of the following special case of a theorem of KOKSMA[5]), which has the advantage that the wanted inequality (3) follows from it immediately:

Let r denote a positive integer; put $P = 2^r$; θ is an irrational number of the type I, described above, so that a number $L = L(\theta)$ exists such that for all integers $q > 1$,

$$|\sin \pi q \theta| > \frac{L}{q^{\eta-1}}.$$

Then if $\varphi(n)$ denotes the polynomial $k u_n$, we have[6])

$$\frac{1}{N}\left|\sum_{M+1}^{M+N} e^{2\pi i \varphi(\nu)}\right| \leqq 50\left(\frac{k^{\eta-1}(r!)^{\eta-1}}{LN}\right)^{\frac{1}{(P-2)(\eta-1)+\frac{P}{2}}}.$$

From this, (3) follows with $\varrho = \eta - 1$ and $\sigma < 1$.

---

[5]) J. F. KOKSMA, Over stelsels Diophantische Ongelijkheden, *Dissertation Groningen*, 1930, Theorem (Stelling) 10, p. 61.

[6]) For the convenience of the reader, this result is obtained by taking the one-dimensional case in KOKSMA's theorem (see [5])) with

$$\Theta = \theta,\ f = \varphi = k u_n,\ g = r!\, k\theta,\ t = 1,\ d = \eta - 1,\ h = kr!$$

and $R = \Delta^r f - g = 0$.

S e c o n d   e x a m p l e. Let $f(t)$ be a p-times differentiable function $(p \geqq 2)$ for $t \geqq 1$, such that $f^{(p)}(t)$ has the same sign for all t, and that

$$\frac{c}{t^{1-\gamma}} \leqq |f^{(p)}(t)| \leqq \frac{C}{t^{1-\gamma}} \qquad (0<\gamma<1,\ 0<c<C),$$

where c, C and γ are independent of t. Then for the sequence $u_n = f(n)$ one has

$$|S(M, N, k)| \leqq A k^{\varrho} N^{\sigma} (M+N)^{\tau} \tag{4}$$

with $\sigma+\tau<1$, $\tau<1/2$, so that theorem II is applicable to the sequence $\{u_n\}$.

The proof of (4) is based on the following lemma of VAN DER CORPUT[7]:

L e m m a. *Let* $M \geqq 0$ $N \geqq 1$, $p \geqq 2$ *be all integers, put* $P = 2^p$ *and let* $g(t)$ *be a real function for* $M \leqq t \leqq M+N$ *which admits a derivative of order* p, *say* $g^{(p)}(t)$ *and suppose that* $g^{(p)}(t) \geqq r$ *for all* t, *or* $g^{(p)}(t) \leqq -r$ *for all* t, *where* r *is independent of* t. *Writing*

$$R = \frac{1}{N} |g^{(p-1)}(M+N) - g^{(p-1)}(M)|$$

one has

$$\left| \sum_{n=M}^{M+N} e^{2\pi i g(n)} \right| \leqq 21 N \left\{ \left(\frac{r}{R^2} \right)^{-\frac{1}{P-2}} + (rN^p)^{-\frac{2}{P}} + \left(\frac{rN}{R} \right)^{-\frac{2}{P}} \right\}. \tag{5}$$

Now apply the lemma to the function $g(t) = kf(t)$, where $f(t)$ satisfies the conditions of our example, and put

$$r = \frac{ck}{(M+N)^{1-\gamma}}, \quad R = \frac{1}{N} \left| \int_M^{M+N} k f^{(p)}(t)\, dt \right|$$

so that

$$R \leqq \frac{1}{N} \int_M^{M+N} \frac{Ck\, dt}{t^{1-\gamma}} \leqq \frac{Ck}{N} \int_0^N t^{\gamma-1} dt = \frac{Ck}{\gamma} \frac{1}{N^{1-\gamma}}.$$

We have now, c_1, c_2, etc. being constants:

$$\left(\frac{r}{R^2} \right)^{-\frac{1}{P-2}} \leqq c_1 k^{\frac{1}{P-2}} (M+N)^{\frac{1-\gamma}{P-2}} N^{-\frac{2(1-\gamma)}{P-2}},$$

$$(rN^p)^{-\frac{2}{P}} \leqq c_2 k^{-\frac{2}{P}} (M+N)^{\frac{2(1-\gamma)}{P}} N^{-\frac{2p}{P}},$$

$$\left(\frac{r}{R} N \right)^{-\frac{2}{P}} \leqq c_3 (M+N)^{\frac{2(1-\gamma)}{P}} N^{-\frac{2}{P}(2-\gamma)}.$$

---

[7]) See e. g. J. G. VAN DER CORPUT, Neue zahlentheoretische Abschätzungen. II, *Math. Zeitschrift*, **29** (1929), pp. 397–426.

Hence, by (5)

$$\left|\sum_{n=M}^{M+N} e^{2\pi i k f(n)}\right| \leqq c_4 k^{\frac{1}{P-2}} (M+N)^{\frac{2(1-\gamma)}{P}} N^{1-\frac{2(1-\gamma)}{P-2}},$$

the inequality being obtained by remarking that, since $0<\gamma<1$, $p \geqq 2$, $P \geqq 4$, one has

$$\frac{2(1-\gamma)}{P} \geqq \frac{1-\gamma}{P-2}$$

and

$$\frac{2(1-\gamma)}{P-2} < \frac{2(2-\gamma)}{P} < \frac{2p}{P}.$$

Writing now $\varrho = \frac{1}{P-2}$, $\sigma = 1 - \frac{2(1-\gamma)}{P-2}$, $\tau = \frac{2(1-\gamma)}{P}$, we remark that, since $P \geqq 4$, $0<\gamma<1$, one has $\tau < 1/2$ and

$$\sigma + \tau = 1 - \frac{2(1-\gamma)}{P-2} + \frac{2(1-\gamma)}{P} < 1$$

so that

$$\left|\sum_{n=M}^{M+N} e^{2\pi i k f(n)}\right| \leqq c_4\, k^{\varrho}\, N^{\sigma} (M+N)^{\tau}$$

with $\sigma+\tau<1$, $\tau<1/2$. We conclude that, under the conditions stated for $f(t)$, Theorem II is applicable to the sequence $u_n = f(n)$.

4. In view of Theorem II the question arises, whether by imposing to the sequence $u_1, u_2, \ldots$ sufficiently strong conditions, e. g. with respect to its discrepancy[8]) $D(N)$, one could avoid any sort of condition on the Fourier coefficients of $f(x)$ and have the relation (1) by merely supposing that the periodic function f belongs to $L^{(2)}$. The answer to this question is negative, as follows from an interesting counterexample due to P. Erdős who communicated it to us verbally: *For every given positive number* $\varepsilon < 1$ *and every decreasing sequence of positive numbers* $\{\delta_n\}$ *for which*

$$\sum_{n=1}^{\infty} \delta_n < \varepsilon \tag{6}$$

a function $f(x)$ *on* $(0, 1)$ *can be constructed, which takes the values* 0 *and* 1 *only, for which* $\int_0^1 f(x)\,dx < \varepsilon$, *whereas the following assertion holds: If* $u_1, u_2, \ldots$ *is any sequence on* $(0, 1)$, *then it can be replaced by a sequence* $u_1', u_2', \ldots$ *such that*

$$|u_n' - u_n| < \delta_n \; (n \geqq 1)$$

[8]) For the definition of discrepancy see e. g. J. F. Koksma, Diophantische Approximationen, *Ergebnisse der Math. und ihrer Grenzgebiete,* IV. 4 (Berlin, 1936), Kap. VIII § 2, p. 90.

whereas for all x

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(u'_n + x) = 1.$$

Now it is obvious that, if the sequence $u_1, u_2, \ldots$ is uniformly distributed (mod 1) with the discrepancy $D(N)$, we can choose $\delta_1, \delta_2, \ldots$ so rapidly decreasing that the sequence $u'_1, u'_2, \ldots$ is also uniformly distributed and has the discrepancy $\leqq 4D(N)$. Therefore:

No matter how fast the positive decreasing function $\varphi(N)$ *may turn to zero as* $N\to\infty$, *if there are sequences* $u_1\, u_2, \ldots$ *for which* $D(N) \leqq \varphi(N)$, *there exist a function* $f(x) \in L^2$ *and certain sequences* $u'_1, u'_2, \ldots$ *satisfying* $D(N) \leqq 4\varphi(N)$, *such that we have*

$$\int_0^1 f(x)\,dx < 1/2 \quad \textit{and} \quad \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(u'_n + x) = 1$$

for every x *on* $(0, 1)$.

We give a complete sketch of the proof. Put without loss of generality $\delta_n = \dfrac{1}{w(n)}$, where $w(1), w(2), \ldots$ denotes an increasing sequence of positive integers. Put $M_1 = 1$, $M_k = (k^2 + k^3 + \ldots + k^{w(k)+1}) M_{k-1}$ $(k \geqq 2)$ and $N_k = w(M_1 + M_2 + \ldots + M_k) + 1$. (Other sequences $M_1, M_2, \ldots$ and $N_1, N_2, \ldots$ would do as well, but it is essential that $M_1, M_2, \ldots$ increase rapidly and $N_1, N_2, \ldots$ still more). Now for $k \geqq 1$ consider in $(0, 1)$ the set T_k consisting of N_k equidistant small segments $\sigma_k^i = \left(\dfrac{i}{N_k}, \dfrac{i}{N_k} + \dfrac{1}{N_k w(k)}\right)$, $i = 0, 1, \ldots, N_k - 1$. Let $f_k(x)$ denote the characteristic function of T_k, whereas $f(x)$ denotes the characteristic function of $T_1 + T_2 + \ldots$. Then

$$f(x) \leqq f_1(x) + f_2(x) + \ldots$$

is a function $\in L^2$ and $\int_0^1 f\,dx < \varepsilon$ by (6).

We now translate the numbers u_n In the first step we move the first M_1 elements of $u_1, u_2, \ldots$ In the second step the following M_2 elements etc.; hence after the k-th step $M_1 + \ldots + M_k$ elements have been moved. In the first step we move u_1 over a distance 0. Now let the $(k-1)$th step be carried out. Then we carry out the k-th step in substeps. In the first substep we remove the first $k^2 M_{k-1}$ elements $(n = M_1 + \ldots + M_{k-1} + 1, \ldots, M_1 + \ldots + M_{k-1} + k^2 M_{k-1})$. In the second step the following $k^3 M_{k-1}$ elements, etc. In the first substep we replace each u_n by an u'_n in such a way that $u'_n + \dfrac{1}{N_k w(k)}$ falls in the lefthand endpoint of a σ_k^i which is nearest to $u_n + \dfrac{1}{N_k w(k)}$ (mod 1).

In the h-th substep (denoted by (k, h)) we replace u_n by an u'_n in such a way that $u'_n + \frac{h}{N_k w(k)}$ falls in the lefthand endpoint of a σ_k^i which is nearest to $u_n + \frac{h}{N_k w(k)}$ (mod 1). Note that (mod 1) each u_n now is moved over a distance $< \frac{1}{N_k} \leqq \delta_k$. Now let x denote an arbitrary real number in (0, 1). Then x for each $k \geqq 2$ lies exactly in one of the $N_k w(k)$ equal parts of length $\frac{1}{N_k w(k)}$ in which we can divide the segment (0, 1), say in the part

$$\frac{h'}{N_k w(k)} \leqq x < \frac{h'+1}{N_k w(k)} \qquad (0 \leqq h' < N_k w(k)).$$

Now there is an uniquely defined integer $h = h(k)$ $(0 \leqq h < w(k))$ such that $h = h'$ (mod $w(k)$).

Consider the elements u_n, which have been moved by the substep (k, h). It is easily proved that the fractional part of the corresponding numbers $u'_n + x$ will belong to one of the segments σ_k^i. Hence $f(u'_n + x) = 1$. Denoting the total number of elements which have been moved after finishing the substep (k, h) by $A(k, h)$ we clearly find

$$\frac{1}{A(k, h)} \sum_{n=1}^{A(k, h)} f(u'_n + x) \geqq \frac{k^{h+1} M_{k-1}}{A(k, h)} \to 1 \text{ as } k \to \infty$$

by the definitions of M_{k-1} and $A(k, h)$. Q. e. d.

(Received July 9, 1949.)

On Pólya frequency functions. II: Variation-diminishing integral operators of the convolution type.

By I. J. SCHOENBERG in Philadelphia, Pa.

1. Introduction and statement of results[1]).

A real matrix $A = \|a_{ik}\|$ $(i=1,\dots m;\ k=1,\dots,n)$ is said to be *totally positive* if all its minors, of any order, are non-negative. In 1930[2]) the author showed that if A is totally positive, then the linear transformation

$$(1) \qquad y_i = \sum_{k=1}^{n} a_{ik} x_k \qquad (i=1,\dots,m)$$

is variation-diminishing in the sense that if $v(x_k)$ denotes the number of variations of sign in the sequence $\{x_k\}$ and $v(y_i)$ the corresponding number in the sequence $\{y_i\}$, then we always have the inequality $v(y_i) \leq v(x_k)$. In the same paper of 1930 the author showed that (1) is certainly variation-diminishing if the matrix A does not possess two minors of equal orders and of opposite signs; also the converse holds to a certain extent: If (1) is variation-diminishing, then A cannot have two minors of equal orders and of opposite signs, *provided the rank of A is $=n$*. The necessary and sufficient conditions in order that (1) be variation-diminishing were found in 1933 by TH. MOTZKIN[3]). Since they will be used in this paper we state them here as follows: Let r be the rank of A; then A should not have two minors of equal orders and of opposite signs if their common order is $<r$, while if their common order is $=r$ then again they should never be of opposite signs if they belong to the same combination of r columns of A.

A function $\Lambda(x)$, $-\infty < x < \infty$, is called a Pólya frequency function (abbreviated P. f. f.) if it satisfies the following three characteristic conditions:

1. $\Lambda(x)$ is measurable.

---

1) A résumé of the results of this paper has appeared under the same title in the *Proceedings of the National Academy of Sciences*, **34** (1948), pp. 164—169.

2) I. J. SCHOENBERG, Über variationsvermindernde lineare Transformationen, *Math. Zeitschrift*, **32** (1930), pp. 321—328.

3) TH. MOTZKIN, *Beiträge zur Theorie der linearen Ungleichungen*, Doctoral dissertation, Basel, 1933 (Jerusalem, 1936), 69 pp., especially Chap. IV.

2. If $x_1 < x_2 < \ldots < x_n$ and $t_1 < t_2 < \ldots < t_n$ then $\det \|\Lambda(x_i - t_j)\|_{1,n} \geqq 0$, $n = 1, 2, 3, \ldots$. For $n = 1$ this means that $\Lambda(x) \geqq 0$.

3. Finally

$$0 < \int_{-\infty}^{\infty} \Lambda(x)\, dx < +\infty.$$

In our previous paper on Pólya frequency functions[4]) it was shown that a P. f. f. $\Lambda(x)$ has a bilateral Laplace transform of the form

$$(2) \qquad \int_{-\infty}^{\infty} e^{-xs} \Lambda(x)\, dx = \frac{1}{\Psi(s)} \qquad (-c < Rs < c \text{ for some } c > 0),$$

where $\Psi(s)$ is an entire function of the type II of Laguerre, Pólya and Schur

$$(3) \qquad \Psi(s) = c e^{-\gamma s^2 - \delta s} \prod_{\nu=1}^{\infty} (1 + \delta_\nu s) e^{-\delta_\nu s} \quad \left(c > 0, \gamma \geqq 0, 0 < \gamma + \sum_1^{\infty} \delta_\nu^2 < \infty\right),$$

and that conversely, the reciprocal of a function $\Psi(s)$ of this type allows of a representation (2) where $\Lambda(x)$ is a P. f. f.

All Pólya frequency functions $\Lambda(x)$ are everywhere continuous with the single exception of the so-called truncated exponential

$$\Lambda(x) = \begin{cases} e^{-x} & \text{if } x \geqq 0, \\ 0 & \text{if } x < 0, \end{cases}$$

and all functions arising from it by a change of scale and origin. Further notable examples of P. f. f. are $\Lambda(x) = e^{-x^2}$, $e^{-|x|}$, e^{x-e^x}, $1/\cosh x$.

Let $f(x)$ be a real function defined for all real x. The number $v(f)$, of variations of sign of $f(x)$ in the range $(-\infty, \infty)$ is defined as follows: If S: $x_1 < x_2 < \ldots < x_n$ is an arbitrary finite increasing sequence of reals, then

$$v(f) = \sup_S v(f(x_i)) \qquad (0 \leqq v(f) \leqq \infty).$$

Let now $L(t)$ be a given real function of bounded variation in the range $-\infty < t < \infty$ which we normalize by the conditions that $L(-\infty) = 0$, $2L(t) = L(t+0) + L(t-0)$; we also rule out the trivial case when $L(t) \equiv 0$. Let us consider the integral transformation

$$(4) \qquad g(x) = \int_{-\infty}^{\infty} f(x-t)\, dL(t),$$

where $f(x)$ is an arbitrary continuous and bounded function. We say that (4)

---

[4]) I. J. Schoenberg, On totally positive functions, Laplace integrals and entire functions of the Laguerre—Pólya—Schur type, *Proceedings of the National Academy of Sciences*, **33** (1947), pp. 11—17. A detailed paper will appear under the title "On Pólya frequency functions. I: Totally positive functions and their Laplace transforms" probably in the *Transactions of the American Mathematical Society*.

is *variation-diminishing* if (4) always implies the inequality

(5) $$v(g) \leqq v(f).$$

Our main result is the following

Theorem 1. *The transformation* (4) *is variation-diminishing if and only if $L(t)$ is either, up to the sign, a cumulative Pólya frequency function*

(6) $$L(t) = \varepsilon \int_{-\infty}^{t} \Lambda(u)\, du,$$

where $\varepsilon = \pm 1$ and $\Lambda(x)$ is a Pólya frequency function, or else $L(t)$ is a step-function with only one jump.

On combining this theorem with (2) and (3) we may restate our result without a distinction of two cases as

Theorem 2. *The transformation* (4) *is variation-diminishing if and only if $L(t)$ has a bilateral Laplace-Stieltjes transform of the form*

(7) $$\int_{-\infty}^{\infty} e^{-st}\, dL(t) = C' e^{\gamma s^2 + \delta s} \prod_{\nu=1}^{\infty} \frac{e^{\delta_\nu s}}{1+\delta_\nu s} \qquad (-c < Rs < c, \text{ for some } c > 0),$$

where $C' \gtrless 0$, $\gamma \geqq 0$, δ, δ_ν real, $\sum_{1}^{\infty} \delta_\nu^2 < \infty$.

It should be noticed that the trivial case when $\gamma = \delta_\nu = 0$, which was excluded in (3), corresponds to a step-function $L(t)$ with only one jump, in which case our transformation (4) becomes

(8) $$g(x) = C' f(x+\delta)$$

and is evidently variation-diminishing.

2. **Proof of the direct part of Theorem 1.**

If (6) holds, the (4) becomes, assuming $\varepsilon = 1$,

(9) $$g(x) = \int_{-\infty}^{\infty} \Lambda(x-t) f(t)\, dt$$

and we are to show that (9) is variation-diminishing within the class of bounded continuous function $f(x)$. This we shall now do even for the wider class of *measurable* and *bounded* functions $f(x)$. Indeed if $v(f) = \infty$ then (5) is trivially verified. Hence we may assume $v(f) < \infty$, $v(f) = m$ say.

Let

(10) $$F(t) = \int_{0}^{t} f(t)\, dt \qquad (-\infty < t < \infty);$$

then (9) becomes

$$g(x)=\int_{-\infty}^{\infty}\Lambda(x-t)\,dF(t). \tag{11}$$

If $x_0<x_1<\ldots<x_n$ we are to show that

$$v(g(x_i))\leqq m, \tag{12}$$

and without losing generality we may assume that $v(g(x_i))=n$, that is

$$g(x_0),\, g(x_1),\ldots,\, g(x_n) \qquad \text{alternate in sign.} \tag{13}$$

Clearly (11) converges uniformly in the $n+1$ points x_i; hence we can find A such that the quantities

$$g_i=\int_{-A}^{A}\Lambda(x_i-t)\,dF(t) \qquad (i=0,1,\ldots,n), \tag{14}$$

also alternate in sign, i. e. $v(g_i)=n$.

But (14) are ordinary Stieltjes integrals. We may therefore divide $(-A, A)$ into N equal parts by $t_0=-A<t_1<\ldots<t_N=A$ such that

$$\bar{g}_i=\sum_{\nu=1}^{N}\Lambda(x_i-t_\nu)\,(F(t_\nu)-F(t_{\nu-1})) \tag{15}$$

are all so close to the respective integrals (14) that we also have

$$v(\bar{g}_i)=n. \tag{16}$$

Since by assumption $v(f)=m$ we see, by (10), that the number of variations of sign in the sequence $F(t_\nu)-F(t_{\nu-1})$ $(\nu=1,\ldots,N)$ is $\leqq m$. But (15) is a variation-diminishing linear transformation in view of the total positivity of the matrix $\|\Lambda(x_i-t_\nu)\|$. By (16) we conclude that $n\leqq m$, hence (12) is established.

3. Proof of the converse part of Theorem I in case when $L(t)$ is continuously differentiable.

Writing $\lambda(t)=L'(t)$ which we assume continuous and clearly also in the Lebesgue class $L(-\infty,\infty)$, we now assume the transformation

$$g(x)=\int_{-\infty}^{\infty}\lambda(x-t)\,f(t)\,dt \tag{17}$$

to be variation-diminishing for bounded continuous $f(t)$ and we are to conclude that $\lambda(x)$ is, up to the sign, a Pólya frequency function. By obvious continuity arguments the variation-diminishing property of (17) immediately generalizes to the case when $f(t)$ is a step-function with a finite number of jumps.

Let

(18) $$x_1 < x_2 < \ldots < x_n; \; t_1 < t_2 < \ldots < t_n;$$

we want to show first that

(19) $$y_i' = \sum_{\nu=1}^{n} \lambda(x_i - t_\nu)\, y_\nu \qquad (i = 1, \ldots, n)$$

is a variation-diminishing linear transformation. To this end, given the y_ν, we define a step-function $f(t)$ as follows: $f(t) = \frac{1}{2h} y_\nu$ if $t_\nu - h \leqq t \leqq t_\nu + h$, $f(t) = 0$ outside the n intervals $[t_\nu - h, t_\nu + h]$ which are assumed not to overlap.

For the values of $g(x_i)$ as given by (17) we now find that

(20) $$g(x_i) = \sum_{\nu=1}^{n} \left(\frac{1}{2h} \int_{t_\nu - h}^{t_\nu + h} \lambda(x_i - t)\, dt \right) y_\nu \qquad (i = 1, \ldots, n).$$

Since $v(f) = v(y_\nu)$, while $v(g(x_i)) \leqq v(g)$, we have

(21) $$v(g(x_i)) \leqq v(y_\nu),$$

showing that (20) is variation-diminishing. Letting $h \to 0$, the continuity of $\lambda(x)$ implies $\lim\limits_{h \to 0} g(x_i) = y_i' = \sum_{\nu=1}^{n} \lambda(x_i - t_\nu)\, y_\nu$, and, by (21), we have

$$v(y_i') \leqq \varliminf_{h \to 0} v(g(x_i)) \leqq v(y_\nu).$$

Thus (19) is indeed variation-diminishing.

This result implies in particular that $\lambda(x)$ never changes sign; for indeed, if it did, we could arrange to have $\lambda(x_1 - t_1)$ and $\lambda(x_2 - t_1)$ of opposite signs and this would contradict the inequality $v(y_i') \leqq v(y_\nu)$ for the values $y_1 = 1$, $y_2 = \ldots = y_n = 0$, if substituted into (19). Thus without loss of generality we may assume to have that

(22) $$\lambda(x) \geqq 0 \quad \text{for all } x, \text{ in particular } \quad \lambda(0) > 0,$$

since the last condition always obtains after a suitable shift of origin.

Let again $\{x_i\}$ *and* $\{t_j\}$ *satisfy* (18). *We wish to prove now that*

(23) $$D_n = \det \| \lambda(x_i - t_j) \|_{1,n} \geqq 0.$$

This is the main point of the proof. We shall establish it first for $n = 2$:

(24) *If* $x_1 < x_2$, $t_1 < t_2$, *then* $D_2 = \begin{vmatrix} \lambda(x_1 - t_1) & \lambda(x_1 - t_2) \\ \lambda(x_2 - t_1) & \lambda(x_2 - t_2) \end{vmatrix} \geqq 0.$

Indeed, suppose we had

(25) $$D_2 < 0.$$

Then let $x_3 = t_3 = \tau > \max(x_2, t_2)$ and consider

$$(26)\qquad D_3 = \begin{vmatrix} \lambda(x_1-t_1) & \lambda(x_1-t_2) & \lambda(x_1-\tau) \\ \lambda(x_2-t_1) & \lambda(x_2-t_2) & \lambda(x_2-\tau) \\ \lambda(\tau-t_1) & \lambda(\tau-t_2) & \lambda(0) \end{vmatrix}.$$

Since $\lambda(x) \geqq 0$ and the integral

$$\int_0^\infty (\lambda(x_1-\tau)+\lambda(x_2-\tau)+\lambda(\tau-t_1)+\lambda(\tau-t_2)\,d\tau$$

converges, we can certainly choose $\tau > \max(x_2, t_2)$ such that each of the four quantities under the integral sign are as small as we please. But then (25), (26) and $\lambda(0) > 0$ clearly imply that

$$D_3 = D_2\,\lambda(0) + \text{ (an arbitrarily small quantity) } < 0$$

for some appropriate value of τ. On the other hand, at least one of the four elements of D_2 is positive, $\lambda(x_1-t_2)$ say, and this implies that

$$(27)\qquad \begin{vmatrix} \lambda(x_1-t_2) & \lambda(x_1-\tau) \\ \lambda(\tau-t_2) & \lambda(0) \end{vmatrix} = \lambda(x_1-t_2)\,\lambda(0) + \text{(small quantity)} > 0.$$

But now we have a contradiction with the properties of variation-diminishing transformation stated in the first paragraph of our introduction: (26) is the determinant of a non-singular variation-diminishing transformation; as such it cannot have two minors, such as (25) and (27), which are of equal orders less than its rank 3, and of opposite signs.

We may now turn to a proof of the general inequality (23). Firstly we recall[5]) that the property (24) means that $\lambda(x)$ is logarithmically concave; this fact and the summability of $\lambda(x)$ imply that

$$(28)\qquad \lim_{x\to\pm\infty} \lambda(x) = 0,$$

even exponentially. Secondly, let $\xi_0, \xi_1, \ldots$ be an infinite sequence of real numbers having the following properties:

1. The sequence $\{\xi_\nu\}$ is monotone increasing.
2. The sequence $\{\xi_\nu\}$ contains every element of the finite sequences $\{x_i\}$ and $\{t_j\}$ appearing in (23).
3. For sufficiently large ν, $\{\xi_\nu\}$ is made up of consecutive integers.

N and r being positive integers let us consider the following three determinants

$$(29)\qquad \begin{aligned} D_{N+1} &= \det\|\lambda(\xi_i-\xi_j)\| && (i,j=0,1,\ldots,N), \\ D_n^* &= \det\|\lambda(\xi_{r\,i}-\xi_{r\,j})\| && (i,j=0,1,\ldots,n-1), \\ D_{n+1}^* &= \det\|\lambda(\xi_{r\,i}-\xi_{r\,j})\| && (i,j=0,1,\ldots,n). \end{aligned}$$

From the property 3 of the sequence $\{\xi_\nu\}$, together with (28), it is clear that

[5]) See the paper mentioned in footnote [4]).

$\lambda(\xi_{ri}-\xi_{rj})\to 0$ as $r\to\infty$, provided that $i\neq j$. By (29) we have $D_n^*\to(\lambda(0))^n$ and $D_{n+1}^*\to(\lambda(0))^{n+1}$; thus we now see that

$$D_n^*>0,\ D_{n+1}^*>0, \tag{30}$$

provided r is chosen large enough. If we now choose $N>rn$, then clearly D_n^* and D_{n+1}^* are minors of D_{N+1}; by the property 2 we may moreover assume also D_n to be a minor of D_{N+1}, by further increasing N, if necessary. We may now prove (23) as follows: We know from the first result of this section that D_{N+1} is the determinant of a variation-diminishing transformation; by (30) the rank of D_{N+1} is $\geqq n+1$. By MOTZKIN's theorem of our introduction we conclude that the two minors D_n and D_n^* cannot have opposite signs, hence $D_n^*>0$ implies that $D_n\geqq 0$. But then $\lambda(x)$ is indeed a Pólya f. f.

4. General proof of the converse part of Theorem 1.

Let $f(t)$ be continuous and bounded and let $g(x)$ be defined by

$$g(x)=\int_{-\infty}^{\infty} f(x-t)\,dL(t). \tag{31}$$

We define two new functions by

$$\lambda(x)=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-(x-t)^2}dL(t) \tag{32}$$

and

$$h(x)=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} g(x-t)\,e^{-t^2}\,dt. \tag{33}$$

Now (31), (32) and (33) imply that

$$h(x)=\int_{-\infty}^{\infty} f(x-u)\,\lambda(u)\,du. \tag{34}$$

Indeed, if we substitute (31) into (33) we find

$$h(x)=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-t^2}dt\int_{-\infty}^{\infty} f(x-t-u)\,dL(u)=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-t^2}dt\int_{-\infty}^{\infty} f(x-u)\,d_uL(u-t).$$

Since $f(t)$ is continuous and bounded and $L(u)$ of bounded variation it is easy to see that we may integrate first under the differential sign d_u obtaining

$$h(x)=\int_{-\infty}^{\infty} f(x-u)\,d_u\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} L(u-t)\,e^{-t^2}dt=\int_{-\infty}^{\infty} f(x-u)\,\frac{d}{du}\left(\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-(u-t)^2}\,L(t)\,dt\right)du.$$

Now

$$\frac{d}{du}\left(\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-(u-t)^2} L(t)\,dt\right) = -\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} 2(u-t)\,e^{-(u-t)^2} L(t)\,dt =$$

$$= -\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} L(t)\,d_t e^{-(u-t)^2} = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-(u-t)^2} dL(t) = \lambda(u)$$

and (34) is established.

By assumption we have $v(g) \leqq v(f)$; but also (33) is a variation-diminishing transformation by the direct part of Theorem 1, hence $v(h) \leqq v(g)$. It follows that $v(h) \leqq v(f)$ and that (34) is therefore a variation-diminishing transformation. Since $\lambda(x)$ is obviously continuous and summable we learn from the previous case already settled, that $\lambda(x)$ is a P. f. f. Therefore

$$\text{(35)} \qquad \int_{-\infty}^{\infty} e^{-sx}\lambda(x)\,dx = \frac{1}{\Psi(s)} = Ce^{\gamma s^2+\delta s}\prod_{\nu=1}^{\infty}\frac{e^{\delta_\nu s}}{1+\delta_\nu s} \qquad (-a < Rs < a),$$

where $C \gtrless 0$, $\gamma \geqq 0$, $0 < \gamma + \Sigma\delta_\nu^2 < \infty$.

Let us consider now the characteristic function of $L(t)$:

$$\text{(36)} \qquad H(it) = \int_{-\infty}^{\infty} e^{-itx}\,dL(x) \qquad (-\infty < t < \infty).$$

The convolution relation (32) now implies the characteristic function relation

$$\frac{1}{\Psi(it)} = e^{-\frac{t^2}{4}} H(it) \qquad (-\infty < t < \infty),$$

or

$$\text{(37)} \qquad H(it) = Ce^{\left(\frac{1}{4}-\gamma\right)t^2+i\delta t}\prod_{\nu=1}^{\infty}\frac{e^{i\delta_\nu t}}{1+i\delta_\nu t} \qquad (-\infty < t < \infty).$$

Passing to norms we have

$$|H(it)|^2 = C^2 e^{2\left(\frac{1}{4}-\gamma\right)t^2} \Big/ \prod_{\nu=1}^{\infty}(1+\delta_\nu^2 t^2),$$

and from this relation we can easily conclude that

$$\text{(38)} \qquad \frac{1}{4} - \gamma \leqq 0.$$

Indeed, let $\varepsilon > 0$ be given and choose n such that $\sum_{n+1}^{\infty}\delta_\nu^2 < \varepsilon$; since $|H(it)|^2$ is bounded, by (36), we have for t^2 sufficiently large the relations

$$C^2 e^{2\left(\frac{1}{4}-\gamma\right)t^2} = |H(it)|^2 \prod_{1}^{n}(1+\delta_\nu^2 t^2)\prod_{n+1}^{\infty}(1+\delta_\nu^2 t^2) < e^{\varepsilon t^2} e^{\varepsilon t^2} = e^{2\varepsilon t^2},$$

hence $\frac{1}{4} - \gamma < \varepsilon$ and thence (38), on letting $\varepsilon \to 0$. With $\gamma_1 = \gamma - \frac{1}{4} \geqq 0$

we may now rewrite (37) as

$$H(it) = Ce^{-\gamma_1 t^2 + i\delta t} \prod_{\nu=1}^{\infty} \frac{e^{i\delta_\nu t}}{1 + i\delta_\nu t} \qquad (-\infty < t < \infty). \tag{39}$$

Two cases may now arise:

1. $\gamma_1 = 0$ and $\delta_1 = \delta_2 = \ldots = 0$. In this case (39) becomes

$$H(it) = \int_{-\infty}^{\infty} e^{-itx} dL(x) = Ce^{i\delta t},$$

hence

$$L(x) = \begin{cases} 0 & \text{if} \quad x < -\delta \\ C & \text{if} \quad x > -\delta. \end{cases}$$

2. $\gamma_1 + \sum_1^{\infty} \delta_\nu^2 > 0$, then $\Psi_1(s) = C^{-1} e^{-\gamma_1 s^2 - \delta s} \prod_{\nu=1}^{\infty} (1 + \delta_\nu s) e^{-\delta_\nu s}$

is of the type of (3); hence its reciprocal may be written as

$$\frac{1}{\Psi_1(s)} = \int_{-\infty}^{\infty} e^{-sx} \varepsilon \Lambda(x)\, dx \qquad (-a < Rs < a),$$

where $\varepsilon = \pm 1$ and $\Lambda(x)$ is a P. f. f. From the last relation, (39) and (36) we now obtain the identity $\int_{-\infty}^{\infty} e^{-itx} dL(x) = \int_{-\infty}^{\infty} e^{-itx} \varepsilon \Lambda(x)\, dx \quad (-\infty < t < \infty)$,

from which we conclude that $L(x) = \varepsilon \int_{-\infty}^{x} \Lambda(u)\, du$.

This concludes the proof of our Theorem 1.

5. The connection of Theorem 1 with a theorem of Pólya.

In 1915 PÓLYA[6]) discovered the following

Theorem 3 (PÓLYA). *Let*

$$\sum_{\nu=0}^{\infty} \frac{(-1)^\nu}{\nu!} \mu_\nu s^\nu \qquad (\mu_0 \neq 0) \tag{40}$$

be a given formal power series with the following property (S)*: If* $f(x)$ *is an arbitrary real polynomial, then the number of real zeros of the polynomial*

$$g(x) = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu}{\nu!} \mu_\nu f^{(\nu)}(x) \tag{41}$$

should never exceed the number of real zeros of the given polynomial $f(x)$. *A power series* (40) *enjoys the property* (S) *if and only if it is the Taylor*

[6]) G. PÓLYA, Algebraische Untersuchungen über ganze Funktionen vom Geschlechte Null und Eins, *Journal für die reine und angewandte Math.*, **145** (1915), pp. 224–249, especially p. 231.

expansion of the reciprocal $1/\Psi(s)$ *of a function of the type* (3), *where, however, the possibility that* $\gamma = \delta_1 = \delta_2 = \ldots = 0$ *is not excluded.*

This theorem is the source of the author's work on "Pólya frequency functions" and also explains the terminology. We now wish to point out that Theorem 3, combined with the representation (2), (3), for the Laplace transforms of Pólya frequency functions, will imply the converse part of our Theorem 1 without much difficulty, provided we make some additional assumption concerning the existence of the moments of $dL(t)$. The following assumption is perhaps stronger than necessary but applies easily: We add to Theorem 1 the additional assumption that

$$(42) \qquad f(s) = \int_{-\infty}^{\infty} e^{-st} dL(t) \qquad \text{converges in a strip } -a < Rs < a.$$

Indeed, (4) is now meaningful if $f(x)$ is a polynomial of degree n, say. Setting

$\mu_\nu = \int_{-\infty}^{\infty} t^\nu dL(t)$, we find that

$$g(x) = \int_{-\infty}^{\infty} f(x-t)\, dL(t) = \int_{-\infty}^{\infty} \sum \frac{(-t)^\nu}{\nu!} f^{(\nu)}(x)\, dx,$$

which may now be written in the form (41). The variation-diminishing property of (4), assumed for bounded functions, is readily shown to hold for polynomials. Once this done, we easily see that the transformation (41) also enjoys the property (S) required by Theorem 3. That $\mu_0 \neq 0$ is seen as follows: Consider $\lambda(x)$ as defined by (32). By assumption $v(\lambda) = 0$. However, $\mu_0 = 0$ would clearly imply $\int_{-\infty}^{\infty} \lambda(x)\, dx = 0$ which contradicts $v(\lambda) = 0$ and the fact that $\lambda(x) \not\equiv 0$ (because $L(t) \not\equiv 0$). By Theorem 3 we now learn that

$$\int_{-\infty}^{\infty} e^{-st} dL(t) = \sum_0^\infty \frac{(-1)^\nu}{\nu!} \mu_\nu s^\nu = \frac{1}{\Psi(s)}.$$

If $\Psi(s)$ reduces to $Ce^{-\delta s}$, then $L(t)$ reduces to a step-function with only one jump. If $\gamma + \Sigma \delta_\nu^2 > 0$, then we have an identity

$$\int_{-\infty}^{\infty} e^{-st} dL(t) = \int_{-\infty}^{\infty} e^{-st} \varepsilon \Lambda(x)\, dx \qquad (-a < Rs < a),$$

where $\varepsilon = \pm 1$ and $\Lambda(x)$ is a P. f. f. This concludes our proof.

(Received July 15, 1949.)

On the Gibbs' phenomenon for Euler means.

By O. SZÁSZ in Cincinnati.

1. Introduction.

Consider a Fourier sine series

$$(1.1)\qquad f(t)\sim\sum_{1}^{\infty} b_n \sin nt, \qquad b_n=\frac{2}{\pi}\int_{0}^{\pi} f(t)\sin nt\,dt,$$

and its partial sums $s_n(t)$ $(n=1,2,3,\ldots)$. FEJÉR proved[1]) that if $f(t)$ is of bounded variation, and if $nt_n\to\tau$ as $t_n\to+0$, then

$$s_n(t_n)\to\frac{2}{\pi}f(+0)\int_{0}^{\tau}\frac{\sin t}{t}dt\equiv\frac{2}{\pi}f(+0)J(\tau), \quad 0\leqq\tau\leqq+\infty.$$

For $\tau=\pi$ $J(\tau)$ attains its maximal value,and

$$\lim_{nt_n\to\pi} s_n(t_n)=\frac{2}{\pi}f(+0)\int_{0}^{\pi}\frac{\sin t}{t}dt=f(+0)\times 1{,}1789797\ldots .$$

Thus the limit-points of the partial sums as $t_n\to 0$ cover an interval which extends beyond $f(+0)$, if $f(+0)\neq 0$. This is called GIBBS' phenomenon, relative to the partial sums.

Our aim is to establish the corresponding phenomenon for Euler means.

For Cesàro means Gibbs' phenomenon was discussed by H. CRAMÉR and T. H. GRONWALL.

2. Euler means of the series $\Sigma\frac{\sin nt}{n}$.

The general Euler means of a sequence $\{s_n\}$ depend on a parameter r, and are defined by the triangular transform

$$\sigma_n(r)=\sigma_n=\sum_{\nu=0}^{n}\binom{n}{\nu}r^\nu(1-r)^{n-\nu}s_\nu, \quad n=0,1,2,\ldots ;$$

We assume $0<r\leqq 1$, in which case the summation method is regular.[2])

[1]) For references see A. ZYGMUND, *Trigonometrical series* (Warszawa, 1935); in particular pp. 179—181.

[2]) For $r=\frac{1}{2}$ it is essentially Euler's series transform; its regularity was first proved by L. D. AMES in 1901. Subsequent results are due to E. JACOBSTHAL, K. KNOPP and R. P. AGNEW.

Nevertheless the Euler means of a Fourier series may be divergent at a point of continuity of the function. Thus we may expect a Gibbs' phenomenon. We first consider the standard series $\sum_1^\infty \frac{\sin nt}{n} = \frac{1}{2}(\pi - t)$, and assume $0 < t < \pi$; now

$$s_0 = 0,\ s_n = \sum_1^n \frac{\sin \nu t}{\nu} = \int_0^t \left(\sum_1^n \cos \nu x\right) dx = -\frac{t}{2} + \int_0^t \frac{\sin (n+\frac{1}{2})x}{2 \sin \frac{1}{2} x} dx,$$

and

$$\sigma_n = -\frac{t}{2} + \int_0^t \frac{1}{2 \sin \frac{1}{2} x} \sum_0^n \binom{n}{\nu} r^\nu (1-r)^{n-\nu} \sin (\nu + \tfrac{1}{2}) x \, dx.$$

The formula

$$\sum_0^n \binom{n}{\nu} r^\nu (1-r)^{n-\nu} e^{i\nu x} = (1 - r + re^{ix})^n$$

now yields

$$\sigma_n + \frac{t}{2} = \frac{1}{2} \Im \int_0^t \frac{1}{\sin \frac{1}{2} x} (1 - r + re^{ix})^n e^{ix/2} dx \text{ [3]}$$

Let $1 - r + re^{ix} = \varrho e^{i\alpha}$, then

(2.1) $\qquad \varrho \cos \alpha = 1 - r + r \cos x, \quad \varrho \sin \alpha = r \sin x,$

(2.2) $\quad \varrho^2 = (1-r)^2 + r^2 + 2r(1-r) \cos x = 1 - 2r(1-r)(1 - \cos x) \leqq 1.$

We assume $0 < x \leqq t \leqq \frac{\pi}{2}$; now

$$\sigma_n + \frac{t}{2} = \frac{1}{2} \int_0^t \frac{1}{\sin \frac{1}{2} x} \varrho^n \sin\left(n\alpha + \frac{x}{2}\right) dx = \frac{1}{2} \int_0^t \cot \frac{x}{2} \varrho^n \sin n\alpha \, dx + \frac{1}{2} \int_0^t \varrho^n \cos n\alpha \, dx.$$

Here

$$\left| \int_0^t \varrho^n \cos n\alpha \, dx \right| < t,$$

hence

$$\int_0^t \varrho^n \cos n\alpha \, dx = \eta t, \quad |\eta| < 1,$$

(2.3)

$$\sigma_n + \frac{1-\eta}{2} t = \frac{1}{2} \int_0^t \varrho^n \cot \frac{x}{2} \sin n\alpha \, dx.$$

We now assume that

(2.4) $\qquad t = t_n, \quad nt_n \to \tau, \ 0 \leqq \tau \leqq \infty, \quad nt_n^2 \to 0.$

We have

$$0 < 1 - \varrho^n = (1 - \varrho) \sum_0^{n-1} \varrho^\nu < n(1 - \varrho),$$

[3]) $\Im$ means the imaginary part.

and from (2.2)

$$1-\varrho^2=4r(1-r)\sin^2 x/2<r(1-r)x^2,$$

so that

(2.5) $$1-\varrho<r(1-r)x^2\leqq x^2/4.$$

It follows that $1-\varrho^n<nx^2/4$, or $1-\varrho^n=\lambda n x^2$, $0<\lambda<1/4$, and

$$\int_0^t \varrho^n\cot\tfrac{1}{2}x\,\sin n\alpha\,dx=\int_0^t\cot\tfrac{1}{2}x\,\sin n\alpha\,dx-n\int_0^t\lambda x^2\cot\tfrac{1}{2}x\,\sin n\alpha\,dx.$$

Now

$$n\left|\int_0^t\lambda x^2\cot\tfrac{1}{2}x\,\sin n\alpha\,dx\right|<n\int_0^t x^2\cot\tfrac{1}{2}x\,dx<\tfrac{1}{3}\pi n t^2=o(1),\text{ as } n\to\infty,$$

in view of (2.4). Thus

(2.6) $$\int_0^t\varrho^n\cot\tfrac{1}{2}x\,\sin n\alpha\,dx=\int_0^t\cot\tfrac{1}{2}x\,\sin n\alpha\,dx+o(1).$$

Next, from (2.2) $\varrho^2\geqq(1-r)^2+r^2\geqq r^2$, $\varrho\geqq r$, and now from (2.1) $r\sin x=$ $=\varrho\sin\alpha\geqq r\sin\alpha$, hence $\alpha<x$. It is well known that $0<x-\sin x<x^3$; now, from (2.1) $\varrho\alpha-rx=\varrho(\alpha-\sin\alpha)-r(x-\sin x)$, hence $|\varrho\alpha-rx|<$ $<\alpha^3+x^3<2x^3$, and

$$|\alpha-rx|\leqq|\varrho\alpha-rx|+(1-\varrho)\alpha<2x^3+x^3=3x^3,$$

or $\alpha=rx+\mu x^3$, $|\mu|<3$. We now have

(2.7) $$\int_0^t\cot\tfrac{1}{2}x\,\sin n\alpha\,dx=\int_0^t\cot\tfrac{1}{2}x\,\sin nrx\,\cos n\mu x^3\,dx+$$
$$+\int_0^t\cot\tfrac{1}{2}x\,\cos nrx\,\sin n\mu x^3\,dx=I_1+I_2$$

and

(2.8) $$|I_2|<3n\int_0^t x^3\cot\tfrac{1}{2}x\,dx=O(nt^3)=o(1).$$

Finally

(2.9) $$I_1=\int_0^t\frac{\sin nrx}{x}dx-\int_0^t\sin nrx\left\{\frac{2}{x}-\frac{\cos\frac{1}{2}x\,\cos n\mu x^3}{\sin\frac{1}{2}x}\right\}dx=T_1(n)+T_2(n),$$

say, where

$$2\sin\tfrac{1}{2}x-x\cos\tfrac{1}{2}x\cos n\mu x^3=2\sin\tfrac{1}{2}x-x\cos\tfrac{1}{2}x+x\cos\tfrac{1}{2}x(1-\cos n\mu x^3).$$

From the mean value theorem $0\leqq2\sin\frac{1}{2}x-x\cos\frac{1}{2}x<\frac{1}{4}x^3$ and

$$\int_0^t\sin nrx\frac{2\sin\frac{1}{2}x-x\cos\frac{1}{2}x}{x\sin\frac{1}{2}x}dx=o\left(\int_0^t x\,dx\right)=o(1).$$

Furthermore

$$\int_0^t \sin nrx \cot \tfrac{1}{2}x\,(1-\cos n\mu x^3)\,dx = O\left(\int_0^t n^2 x^5\,dx\right) = O(n^2t^6) = o(1),$$

so that $T_2(n) \to 0$.

Collecting (2.3), (2.6), (2.7), (2.8) and (2.9), we find

$$2\sigma_n + (1-\eta)t_n = 2\int_0^{t_n} \frac{\sin nrx}{x}\,dx + o(1) = 2\int_0^{nrt_n} \frac{\sin y}{y}\,dy + o(1),$$

or

$$\sigma_n \to \int_0^{r\tau} \frac{\sin y}{y}\,dy, \text{ as } nt_n \to \tau, \quad 0 \leqq \tau \leqq \infty, \quad \text{and } nt_n^2 \to 0.$$

The complete discussion of the case $\limsup nt_n^2 > 0$ is more complicated, but we can prove in any case that $\limsup \sigma_n(t_n) \leqq \int_0^{\pi} \frac{\sin t}{t}\,dt$. This follows from

$$\limsup \sigma_n(t_n) \leqq \limsup s_n(t_n) \sum_0^n \binom{n}{\nu} r^\nu (1-r)^{n-\nu} = \int_0^{\pi} \frac{\sin t}{t}\,dt.$$

Summarizing, we have proved the following theorem:

Theorem 1. *For the Euler means of the series* $\sum_1^\infty \frac{\sin nt}{n}$ *we have*

$$\lim \sigma_n(t_n) = \int_0^{r\tau} \frac{\sin y}{y}\,dy \quad as \quad nt_n \to \tau, \quad nt^2 \to 0,$$

and always

$$\limsup \sigma_n(t_n) \leqq \int_0^{\pi} \frac{\sin t}{t}\,dt.$$

3. Gibbs' phenomenon of Euler means for a class of Fourier series.

Suppose that the series (1.1) has a simple discontinuity at the point 0: $f(+0) = \pi A/2 > 0$, and let

(3.1) $\quad \varphi(t) = f(t) - A(\pi - t)/2 \sim \Sigma(b_n - An^{-1}) \sin nt \equiv \Sigma \beta_n \sin nt,$

so that $\varphi(t)$ is continuous at $t=0$. Suppose further that the series (3.1) is uniformly summable by Euler means at $t=0$. The behaviour of $\sigma_n\{f, t_n\}$ is then the same as for the series $A\sum' \frac{\sin nt}{n}$. In particular, if the series (3.1) is uniformly *convergent* at $t=0$, then its Euler means present the Gibbs' phenomenon exhibited in Theorem 1. A case in point is, when [4])

$$\lim_{\lambda \downarrow 1} \limsup_{n\to\infty} \sum_{n}^{\lambda n} (|b_\nu| - b_\nu) = 0. \tag{3.2}$$

Thus we have the theorem:

Theorem 2. *If* (3.2) *holds, and if $f(t)$ has a simple discontinuity at $t=0$, then the Euler means of the series* (1.1) *present the same Gibbs' phenomenon as in Theorem* 1.

Final remark. We can discuss in a similar manner the Gibbs' phenomenon for certain Hausdorff means and for Borel summability.

UNIVERSITY OF CINCINNATI,
NATIONAL BUREAU OF STANDARDS.

(Received August 13, 1949.)

[4]) O. Szász, On uniform convergence of Fourier series, *Bulletin American Math. Society*, **50** (1949), pp. 587—595; in particular Theorem 2.

Extremum problems for non-negative sine polynomials.

By
W. W. ROGOSINSKI and G. SZEGŐ
in Newcastle upon Tyne (England) in Stanford (California)

In various chapters of the theory of Fourier series and elsewhere non-negative trigonometrical polynomials

$$(0.1)\qquad T(\vartheta)\equiv\frac{1}{2}a_0+(a_1\cos\vartheta+b_1\sin\vartheta)+\ldots+(a_n\cos n\vartheta+b_n\sin n\vartheta)$$

play an important rôle. For instance, the non-negative character of the arithmetic means of the polynomials

$$(0.2)\qquad \frac{1}{2}+\cos\vartheta+\ldots+\cos n\vartheta$$

is the basic fact in FEJÉR's theory of summability of Fourier series. Similarly, certain sine polynomials, non-negative for $0\leqq\vartheta\leqq\pi$ (in the range $\langle 0,\pi\rangle$), are frequently of importance. As an example we quote GRONWALL's polynomials

$$(0.3)\qquad \sin\vartheta+\frac{1}{2}\sin 2\vartheta+\ldots+\frac{1}{n}\sin n\vartheta.$$

In 1915, L. FEJÉR and F. RIESZ [2][1]) gave a parametric representation of fundamental importance for non-negative trigonometrical polynomials. By means of this representation, L. FEJÉR and others determined in

$$(0.4)\qquad T(\vartheta)\leqq\frac{1}{2}a_0(n+1);\quad a_k^2+b_k^2\leqq a_0^2\cos^2\pi\Big/\left(\left[\frac{n}{k}\right]+2\right)$$

the maxima for such polynomials and for their coefficients, when the constant term $\frac{1}{2}a_0$ and the degree n are prescribed. It should be noted that FEJÉR's problem remains essentially the same if the subclass of non-negative cosine polynomials is considered[2]).

1) The numbers refer to the list of references at the end of this paper.

2) Cf. L. FEJÉR [3] where a survey of the relevant literature can be found. An extension of these results to 'finite' Fourier integrals (which are in a certain sense the analogues of trigonometrical polynomials) has been given more recently by BOAS and KAC [1].

A completely new situation arises if one considers sine polynomials

(0.5) $$S(\vartheta) \equiv b_1 \sin\vartheta + b_2 \sin 2\vartheta + \ldots + b_n \sin n\vartheta \qquad (b_n \neq 0)$$

of given degree n which are non-negative in the range $\langle 0, \pi \rangle$. It is the class of these polynomials we discuss in the present paper. Clearly $b_1 \geqq 0$ and $b_1 = 0$ is only possible when $S(\vartheta)$ vanishes identically. We shall usually normalise by assuming that $b_1 = 1$.

First we determine the maximum of $S(\vartheta)$ for fixed ϑ in $\langle 0, \pi \rangle$, and find:

$$(0.6)_o \quad S(\vartheta) \leqq \operatorname{Max} \begin{cases} \dfrac{1}{4\sin^2\vartheta} \{(n+2)\sin\vartheta - \sin(n+2)\vartheta\} \\ \dfrac{1}{4\sin^3\vartheta} \displaystyle\sum_0^{(n-3)/2} \dfrac{\{(k+3)\sin(k+1)\vartheta - (k+1)\sin(k+3)\vartheta\}^2}{(k+1)(k+3)} \end{cases};$$

$$(0.6)_e \quad S(\vartheta) \leqq \operatorname{Max} \begin{cases} \dfrac{1}{2} \dfrac{\cot\frac{1}{2}\vartheta}{\sin^2\vartheta} \displaystyle\sum_0^{(n-2)/2} \dfrac{\{(k+2)\sin(k+1)\vartheta - (k+1)\sin(k+2)\vartheta\}^2}{(k+1)(k+2)} \\ \dfrac{1}{2} \dfrac{\tan\frac{1}{2}\vartheta}{\sin^2\vartheta} \displaystyle\sum_0^{(n-2)/2} \dfrac{\{(k+2)\sin(k+1)\vartheta + (k+1)\sin(k+2)\vartheta\}^2}{(k+1)(k+2)} \end{cases},$$

when n is odd or even, respectively. In particular, when $\vartheta = 0$,

$$1 + 2b_2 + 3b_3 + \ldots + nb_n \leqq \begin{cases} (n+1)(n+2)(n+3)/24 & (n \text{ odd}) \\ n(n+2)(n+4)/24 & (n \text{ even}). \end{cases}$$

The determination of the maxima and minima for the coefficients b_k is rather involved[3]). We have computed them in the cases b_2, b_3 and b_{n-1}, b_n. In other cases, in particular for b_4 and b_5, we discuss relevant methods of determination. Our main results are:

(0.7) $$|b_2| \leqq \begin{cases} 2\cos 2\pi/(n+3) & (n \text{ odd}) \\ 2\cos\vartheta_0 & (n \text{ even}), \end{cases}$$

where ϑ_0 is the least positive root of

(0.8) $$(n+4)\sin(n+2)\vartheta/2 + (n+2)\sin(n+4)\vartheta/2 = 0.$$

Next,

$(0.9)_e$ $$1 - 2\cos\pi/(n'+3) \leqq b_3 \leqq 1 + 2\cos 2\pi/(n'+3),$$

$(0.9)_o$ $$1 - 2\cos\vartheta_1 \leqq b_3 \leqq 1 + 2\cos 2\pi/(n'+3),$$

according to whether $n' = [\frac{1}{2}(n-1)]$ is even or odd, respectively. Here ϑ_1 is the least positive root of

(0.10) $$(n'+4)\cos(n'+2)\vartheta/2 + (n'+2)\cos(n'+4)\vartheta/2 = 0.$$

---

[3]) The estimate

$$|b_k| = \frac{2}{\pi} \left| \int_0^\pi S(\vartheta) \sin\vartheta \frac{\sin k\vartheta}{\sin\vartheta} d\vartheta \right| \leqq k b_1 = k$$

is trivial.

Further

(0. 11) $$|b_{n-1}| \leq 1, \quad -(n-2)/(n+2) \leq b_{n-1} \leq 1,$$

according to n being odd or even, respectively. Finally, in the same two cases,

(0. 12) $$-(n-1)/(n+3) \leq b_n \leq 1, \quad |b_n| \leq n/(n+2).$$

The introductory § 1 contains general remarks concerning various methods dealing with problems of our kind. In § 2 we determine the maximum of $S(\vartheta)$, in § 3 the extrema for b_{n-1} and b_n, in § 4 for b_2, and in § 5 for the less simple case b_3. In § 6 some formal properties of orthogonal polynomials are discussed which are useful in dealing with the general b_k. The last § 7 deals, in particular, with b_4 and b_5.

§ 1. General remarks.

1. For given degree n and given $b_1 (=1)$ we put

(1. 1) $$\underline{B}(k, n) = \operatorname{Min} b_k, \quad \overline{B}(k, n) = \operatorname{Max} b_k.$$

Now, if the sine polynomial $S(\vartheta)$ is positive in $\langle 0, \pi \rangle$, then so is $S(\pi-\vartheta) = \Sigma(-1)^{k-1} b_k \sin k\vartheta$. Hence

(1. 2) $$\underline{B}(k, n) = -\overline{B}(k, n) \qquad (k \text{ even}).$$

Also

(1. 3) $$S^*(\vartheta) = \frac{1}{2}\{S(\vartheta) + S(\pi-\vartheta)\} = b_1 \sin\vartheta + b_3 \sin 3\vartheta + \dots$$

is non-negative. If n is even, then S^* is of degree at most $n-1$. It follows that

(1. 4) $$\underline{B}(k, n) = \underline{B}(k, n-1), \quad \overline{B}(k, n) = \overline{B}(k, n-1) \qquad (k \text{ odd}, n \text{ even}).$$

2 Let $0 \leq \vartheta \leq \pi$. Any non-negative sine polynomial can be written in the form

(1. 5) $$S(\vartheta) = \sin\vartheta \sum_1^n b_k \sin k\vartheta / \sin\vartheta = \sin\vartheta \; P(\cos\vartheta)$$

where

(1. 6) $$P(\cos\vartheta) = \frac{1}{2} a_0 + a_1 \cos\vartheta + \dots + a_{n-1} \cos(n-1)\vartheta$$

is a non-negative cosine polynomial of degree $n-1$; the converse is also true. Here

(1. 7) $$2 b_k = a_{k-1} - a_{k+1}$$

where $a_n = a_{n+1} = 0$. Now, according to L. FEJÉR and F. RIESZ, any non-negative trigonometrical polynomial $P(\cos\vartheta)$ admits of the parametric representation

(1. 8) $$P(\cos\vartheta) = |c_0 + c_1 e^{i\vartheta} + \dots + c_{n-1} e^{i(n-1)\vartheta}|^2,$$

where the c_ν are (arbitrary) real constants. Hence, by (1.7), $b_k = \Phi_k(c_0, c_1, \ldots, c_{n-1})$ is a certain quadratic form of the c_ν. In particular,

$$(1.9)\quad b_1 = \Phi_1 = \frac{1}{2}(a_0 - a_2) = c_0^2 + c_1^2 + \ldots + c_{n-1}^2 - (c_0 c_2 + c_1 c_3 + \ldots + c_{n-3} c_{n-1})$$

is positive definite.

There are then, theoretically, two possibilities of computing $\underline{B}(k, n)$ and $\overline{B}(k, n)$:

(i) We can either form the characteristic equation

$$(1.10)\qquad |\Phi_k - \lambda \Phi_1| = 0$$

and obtain our quantities as the least and greatest roots of this equation.

(ii) Alternatively, we may form the system of linear equations in the c_ν corresponding to (1.10) and solve this system. This method works satisfactorily in the cases b_{n-1} and b_n.

In general, however, the method based on (1.10) is not easily adaptable for obtaining explicit results, in particular when n is large.

3. We prefer to base our actual discussion on the following theorem of LUKÁCS [**4,** pp. 4—5][4]):

Any polynomial $P(x)$ of degree $\widehat{P} = N$, which is non-negative in $\langle -1, 1 \rangle$, can be represented in the form

$$(1.11)_e\qquad P(x) = A^2(x) + (1 - x^2) B^2(x);\ \widehat{A} \leqq \frac{N}{2},\ \widehat{B} \leqq \frac{N-2}{2},$$

$$(1.11)_o\qquad P(x) = (1 - x) C^2(x) + (1 + x) D^2(x);\ \widehat{C} \leqq \frac{N-1}{2},\ \widehat{D} \leqq \frac{N-1}{2},$$

according to N being even or odd.

Now, by (1.5), if we put $x = \cos\vartheta$,

$$(1.12)\qquad S(\vartheta) = \sin\vartheta\, P(x) = (1 - x^2)^{1/2} P(x),$$

where $P(x)$ is a non-negative polynomial of degree $N = n - 1$ in $\langle -1, 1 \rangle$. It is clear that, when N is even, say, we can restrict P to range over polynomials of the type A^2 or $(1 - x^2) B^2$. Max $S(\vartheta)$ is then the greater (not smaller) maximum obtained in the two cases. A similar remark applies to $\underline{B}(k, n)$ and $\overline{B}(k, n)$, and to the case when N is odd.

4. We have

$$(1.13)\qquad P(x) = \sum_1^n b_k \sin k\vartheta / \sin\vartheta = \sum_1^n b_k U_{k-1}(x).$$

Here

$$(1.14)\quad U_k(x) = U_k(\cos\vartheta) = \sin(k + 1)\vartheta / \sin\vartheta = 2^k x^k + A_k x^{k-2} + \ldots$$

is the familiar Tchebychev polynomial of the second kind. The polynomials

---

[4]) This theorem can also be derived from the results of FEJÉR and RIESZ.

$\sqrt{\frac{2}{\pi}}\,U_k$ form an orthonormal system with the weight function $(1-x^2)^{1/2}$ over the range $\langle -1, 1\rangle$. It follows that

$$(1.15)\qquad \int_{-1}^{1} x^k U_k(x)\,(1-x^2)^{1/2}\,dx = \pi 2^{-(k+1)}.$$

Also

$$(1.16)\qquad b_k = \frac{2}{\pi}\int_{-1}^{1} P(x)\,U_{k-1}(x)\,(1-x^2)^{1/2}\,dx.$$

We shall also require orthogonal polynomials over $\langle -1, 1\rangle$ corresponding to the weight functions $w(x) = (1-x^2)^{3/2}$ and $w(x) = (1-x)(1-x^2)^{1/2}$. The former are[5])

$$(1.17)\qquad V_k(x) = (x^2-1)^{-1}\left[\frac{U_{k+2}(x)}{k+3} - \frac{U_k(x)}{k+1}\right] = \frac{2^{k+2}}{k+3}x^k + \dots,$$

so that, by (1.15),

$$(1.18)\quad \int_{-1}^{1} V_k^2(x)\,(1-x^2)^{3/2}\,dx = \int_{-1}^{1}\left(\frac{U_k(x)}{k+1} - \frac{U_{k+2}(x)}{k+3}\right)\left(\frac{2^{k+2}}{k+3}x^k + \dots\right)(1-x^2)^{1/2}dx$$
$$= \frac{2\pi}{(k+1)(k+3)}.$$

Similarly, when $w(x) = (1-x)(1-x^2)^{1/2}$, we have the orthogonal polynomials

$$(1.19)\qquad W_k(x) = (x-1)^{-1}\left(\frac{U_{k+1}(x)}{k+2} - \frac{U_k(x)}{k+1}\right) = \frac{2^{k+1}}{k+2}x^k + \dots$$

with

$$(1.20)\qquad \int_{-1}^{1} W_k^2(x)\,(1-x)\,(1-x^2)^{1/2}\,dx = \pi/(k+1)(k+2).$$

5. Our problem is of the general type of determining the extrema of a quotient

$$(1.21)\qquad \int_{\alpha}^{\beta} u^2(x)\,h(x)\,w(x)\,dx \;:\; \int_{\alpha}^{\beta} u^2(x)\,w(x)\,dx,$$

where $w(x)$ is a given weight function and $h(x)$ a given polynomial; $u(x)$ is an arbitrary polynomial of given degree whose coefficients vary through all real values not all zero[6]).

In the cases b_2 and b_3 we shall have $h(x) = x$. This is the so called '*problem of the centroid*', first treated by Tchebychev.

---

[5]) Compare (6.6).

[6]) Actually, we shall have either $\alpha = -1$, $\beta = 1$ or $\alpha = 0$, $\beta = 1$.

Its solution is as follows: [Cf. 4, Theorem 7.72.1, p. 183; we follow (apart from slight changes) the notation of **4.**]

Let $w(x)$ be a given weight function over $\langle\alpha, \beta\rangle$ and the $p_k(x)$ be the orthonormal polynomials associated with it. Let $f(x)$ run through all polynomials of given degree N and non-negative in $\langle\alpha, \beta\rangle$. Finally, let $\overline{M}$ and $\underline{M}$ be the maximum and minimum of the quotient

$$\int_\alpha^\beta f(x)\, x\, w(x)\, dx \;:\; \int_\alpha^\beta f(x)\, w(x)\, dx. \tag{1.22}$$

If $N=2m$, then $\overline{M}$ is the greatest and $\underline{M}$ is the least zero of $p_{m+1}(x)$. If $N=2m+1$, then $\overline{M}$ is the greatest zero of $p_{m+2}(\alpha)\, p_{m+1}(x)-p_{m+1}(\alpha)\, p_{m+2}(x)$, and $\underline{M}$ is the least zero of $p_{m+2}(\beta)\, p_{m+1}(x)-p_{m+1}(\beta)\, p_{m+2}(x)$.

We note that extremum problems of our type are normally treated by the Gauss-Jacobi method of mechanical quadrature. We use instead, in §§ 6 and 7, certain formal identities for orthogonal polynomials associated with $w(x)$ and $h(x)\, w(x)$.

§ 2. The maximum of $S(\vartheta)$.

1. First, let n be odd, so that the degree $N=n-1$ of $P(x)$, in (1.13), is even. By $(1.11)_e$, we may assume $P=A^2$ or $P=(1-x^2)B^2$. The maximum of $S(\vartheta)$ is then the greater of the two maxima obtained in each case.

(i) Let $P=A^2$ and $A(x)=\sum_0^h \alpha_k U_k(x)$, where $h=\frac{1}{2}(n-1)$. Since the $\sqrt{\frac{2}{\pi}}\, U_k$ are orthonormal with weight $(1-x^2)^{1/2}$, we have, by (1.16),

$$b_1=\frac{2}{\pi}\int_{-1}^{1} A^2(x)\,(1-x^2)^{1/2}\, dx=\sum_0^h \alpha_k^2=1. \tag{2.1}$$

Hence, by Cauchy's inequality,

$$\begin{aligned} P(x) &\leq \sum \alpha_k^2 \cdot \sum U_k^2(x)=\sum U_k^2(x)= \\ &=\sum\left(\frac{\sin(k+1)\vartheta}{\sin\vartheta}\right)^2=\sum_0^h \frac{1-\cos 2(k+1)\vartheta}{2\sin^2\vartheta}= \\ &=\frac{1}{2\sin^2\vartheta}\left[\frac{n+1}{2}+\frac{1}{2}-\frac{\sin(n+2)\vartheta}{2\sin\vartheta}\right], \end{aligned} \tag{2.2}$$

so that by (1.12)

$$S(\vartheta)\leq\frac{1}{4\sin^2\vartheta}\{(n+2)\sin\vartheta-\sin(n+2)\vartheta\}, \tag{2.3}$$

which is the first inequality $(0.6)_o$.

Clearly, equality in (2.2) and hence in (2.3) can be attained.

(ii) Let $P=(1-x^2)B^2$ and $B(x)=\sum_0^h \beta_k V_k(x)$ where $h=\frac{1}{2}(n-3)$. Then, by (1.18),

$$(2.4)\qquad b_1=\frac{2}{\pi}\int_{-1}^{1} B^2(x)(1-x^2)^{3/2}dx=4\sum_0^h \frac{\beta_k^2}{(k+1)(k+3)}=1.$$

Hence

$$(2.5)\quad B^2(x)\leqq\sum\frac{\beta_k^2}{(k+1)(k+3)}\cdot\sum(k+1)(k+3)V_k^2(x)=\frac{1}{4}\sum_0^h(k+1)(k+3)V_k^2(x)$$

which is equivalent to the second inequality $(0.6)_o$.

2. If n is even we have the two cases $P(x)=(1\pm x)C^2(x)$. It suffices to consider the case of the factor $1-x$, the two cases changing into each other on replacing x by $-x$, that is ϑ by $\pi-\vartheta$.

Putting $C(x)=\sum_0^h \gamma_k W_k(x)$, where $h=\frac{1}{2}(n-2)$, we have, by (1.20),

$$(2.6)\qquad b_1=\frac{2}{\pi}\int_{-1}^{1} C^2(x)(1-x)(1-x^2)^{1/2}dx=2\sum_0^h\frac{\gamma_k^2}{(k+1)(k+2)}=1.$$

Hence

$$(2.7)\qquad C^2(x)\leqq\frac{1}{2}\sum_0^h(k+1)(k+2)W_k^2(x),$$

which is equivalent to the first inequality $(0.6)_e$. The second is obtained on changing ϑ into $\pi-\vartheta$.

§ 3. The extrema of b_n and b_{n-1}.

1. Let n be odd. Again we have two cases:

(i) Let $P(x)=A^2(x)$ where $A(x)=\sum_0^h \alpha_k U_k(x)$ and $h=\frac{1}{2}(n-1)$. By (1.14), $P(x)=\alpha_h^2 2^{2h}x^{2h}+\ldots$ Hence, using (1.15) and (1.16),

$$(3.1)\quad b_n=\frac{2}{\pi}\int_{-1}^{1}P(x)U_{n-1}(x)(1-x^2)^{1/2}dx=\frac{2}{\pi}\alpha_h^2 2^{2h}\int_{-1}^{1}x^{2h}U_{2h}(x)(1-x^2)^{1/2}dx=$$
$$=\frac{2}{\pi}\alpha_h^2 2^{2h}\pi 2^{-(2h+1)}=\alpha_h^2\leqq 1,$$

by (2.1). Also $b_n\geqq 0$.

(ii) Let $P(x)=(1-x^2)B^2(x)=(1-x^2)\left\{\sum_0^h\beta_k V_k(x)\right\}^2$ where $h=\frac{1}{2}(n-3)$. Then, by (1.17),

$$(3.2)\qquad P(x)=-\beta_h^2\frac{2^{2(h+2)}}{(h+3)^2}x^{2h+2}+\ldots$$

Hence, as in (3. 1), since $n-1=2h+2$,

$$b_n = -\frac{2}{\pi}\beta_h^2 \frac{2^{2(h+2)}}{(h+3)^2}\pi 2^{-(2h+3)} = -\frac{4\beta_h^2}{(h+3)^2} \geqq -\frac{(h+1)(h+3)}{(h+3)^2} = -\frac{n-1}{n+3}, \tag{3.3}$$

by (2. 4). Also $b_n \leqq 0$. This establishes (0. 12), when n is odd.

2. If n is even, we may take

$$P(x) = (1-x)\,C^2(x) = (1-x)\left\{\sum_0^h \gamma_k W_k(x)\right\}^2 \quad \text{where } h = \frac{1}{2}(n-2).$$

By (1. 19),

$$P(x) = -\gamma_h^2 \frac{2^{2(h+1)}}{(h+2)^2} x^{2h+1} + \dots, \tag{3.4}$$

and we find, as above, using (2. 6),

$$b_n = -\frac{2}{\pi}\gamma_h^2 \frac{2^{2(h+1)}}{(h+2)^2}\pi 2^{-2(h+1)} = -\frac{2\gamma_h^2}{(h+2)^2} \geqq -\frac{(h+1)(h+2)}{(h+2)^2} = -\frac{n}{n+2}, \tag{3.5}$$

which establishes (0. 12), when n is even.

3. For b_{n-1} we may assume that n is odd, since for even n the case reduces, by (1. 4), to that of the last coefficient. Again, we have our two cases.

(i) We take

$$\begin{aligned} P(x) &= \left(\sum_0^h \alpha_k U_k(x)\right)^2 = (\alpha_h 2^h x^h + \alpha_{h-1} 2^{h-1} x^{h-1} + \dots)^2 = \\ &= 2^{2h}\alpha_h^2 x^{2h} + 2^{2h}\alpha_h \alpha_{h-1} x^{2h-1} + \dots;\ h = \frac{1}{2}(n-1). \end{aligned} \tag{3.6}$$

Hence, by (1. 16) and (2. 1),

$$\begin{aligned} b_{n-1} &= \frac{2}{\pi}\int_{-1}^{1} P(x)\, U_{2h-1}(x)\,(1-x^2)^{1/2}\,dx = \\ &= \frac{2}{\pi} 2^{2h}\alpha_h \alpha_{h-1}\pi 2^{-2h} = 2\alpha_h\alpha_{h-1} \leqq \alpha_h^2 + \alpha_{h-1}^2 \leqq 1. \end{aligned} \tag{3.7}$$

(ii) We take

$$\begin{aligned} P(x) &= (1-x^2)\left(\sum_0^h \beta_k V_k(x)\right)^2 = (1-x^2)\left[\frac{2^{h+2}}{h+3}\beta_h x^h + \frac{2^{h+1}}{h+2}\beta_{h-1}x^{h-1} + \dots\right]^2 = \\ &= -\frac{2^{2(h+2)}}{(h+3)^2}\beta_h^2 x^{2h+2} - \frac{2^{2(h+2)}\beta_h\beta_{h-1}}{(h+3)(h+2)} x^{2h+1} + \dots;\ h = \frac{n-3}{2}. \end{aligned} \tag{3.8}$$

Hence, by (2. 4),

$$\begin{aligned} b_{n-1} &= -\frac{2}{\pi}\frac{2^{2(h+2)}\beta_h\beta_{h-1}}{(h+3)(h+2)}\cdot \pi 2^{-2h} = -\frac{8\beta_h\beta_{h-1}}{(h+3)(h+2)} \geqq \\ &\geqq -4\frac{\beta_{h-1}^2 + \beta_h^2}{(h+3)(h+2)} > -4\left[\frac{\beta_{h-1}^2}{h(h+2)} + \frac{\beta_h^2}{(h+1)(h+3)}\right] \geqq -1. \end{aligned} \tag{3.9}$$

This completes the proof of (0. 11) when n is odd. When n is even, (0. 11) follows from the first formula (0. 12) on replacing n by $n-1$.

§ 4. The extrema for b_2.

By (1.2), it suffices to determine $\overline{B}(2,n)$. Since $U_1(x)=2x$ we have

$$(4.1)\qquad b_2:b_1=2\int_{-1}^{1}xP(x)(1-x^2)^{1/2}dx:\int_{-1}^{1}P(x)(1-x^2)^{1/2}dx,$$

which is a special case of the problem of the centroid (1.2). If n is odd, $\widehat{P}=n-1$ is even, and $\overline{B}(2,n)$ is twice the greatest zero of

$$U_{\frac{n+1}{2}}(x)=\frac{\sin\frac{n+3}{2}\vartheta}{\sin\vartheta},\quad \text{i. e. } \overline{B}(2,n)=2\cos\frac{2\pi}{n+3}.$$

If n is even, then $\overline{B}(2,n)=2\cos\vartheta_0$, where $x_0=\cos\vartheta_0$ is the greatest root of

$$(4.2)\qquad U_{\frac{n+2}{2}}(-1)\,U_{\frac{n}{2}}(x)-U_{\frac{n}{2}}(-1)\,U_{\frac{n+2}{2}}(x)=0,$$

which is equivalent to (0.8).

§ 5. The extrema for b_3.

By (1.3) and (1.4) we may assume that n is odd and that $P(x)=Q(x^2)$ where $\widehat{Q}=n'=\frac{1}{2}(n-1)$. Since $U_2(x)=4x^2-1$ we have

$$(5.1)\qquad b_3:b_1=\int_0^1(4t-1)\,Q(t)(1-t)^{1/2}t^{-1/2}dt:\int_0^1 Q(t)(1-t)^{1/2}t^{-1/2}dt=4T-1,$$

say. Thus our problem is again a special case of the problem of the centroid.

Now the $p_k(t)=U_{2k}(\sqrt{t})$ are, plainly, orthogonal polynomials over $\langle 0,1\rangle$ associated with the weight function $(1-t)^{1/2}t^{1/2}$.

If $n'=2m$, then $p_{m+1}(t)=U_{n'+2}(\sqrt{t})$ and hence

$$(5.2)\qquad \operatorname{Max} T=\cos^2\frac{\pi}{n'+3},\quad \operatorname{Min} T=\cos^2\left(\frac{\pi}{2}\,\frac{n'+2}{n'+3}\right)=\sin^2\frac{\pi}{2(n'+3)},$$

$$(5.3)\qquad \overline{B}(3,n)=1+2\cos\frac{2\pi}{n'+3},\quad \underline{B}(3,n)=1-2\cos\frac{\pi}{n'+3}.$$

If $n'=2m+1$, we need the greatest zero of

$$(5.4)\qquad U_{2m+4}(0)\,U_{2m+2}(\sqrt{t})-U_{2m+2}(0)\,U_{2m+4}(\sqrt{t})=$$
$$=(-1)^{m+2}\frac{\sin(2m+3)\vartheta+\sin(2m+5)\vartheta}{\sin\vartheta}=(-1)^{m+2}\frac{2\sin(3m+4)\vartheta\cos\vartheta}{\sin\vartheta},$$

where $t=\cos^2\vartheta$, which leads to the right half of $(0.9)_0$; the least zero of

$$(5.5)\qquad U_{2m+4}(1)\,U_{2m+2}(\sqrt{t})-U_{2m+2}(1)\,U_{2m+4}(\sqrt{t})=$$
$$=\frac{(2m+5)\sin(2m+3)\vartheta-(2m+3)\sin(2m+5)\vartheta}{\sin\vartheta}$$

similarly gives the left half of $(0.9)_0$.

§ 6. Identities involving orthogonal polynomials

1. Let $w(x)$ be a weight function over $\langle\alpha, \beta\rangle$, and let the $p_m(x) = k_m x^m + \dots$, where $k_m > 0$, be the associated orthonormal polynomials. We introduce the moments

(6. 1) $$c_m = \int_\alpha^\beta x^m w(x)\, dx$$

The determinants $D_m = [c_{p+q}]_0^m$ are then positive, and we have, for $m \geq 1$,[7])

(6. 2) $$p_m(x) = (D_{m-1} D_m)^{-1/2} [c_{p+q} x - c_{p+q+1}]_0^{m-1},$$

(6. 3) $$k_0 = D_0^{-1/2}, \quad k_m = (D_{m-1}/D_m)^{1/2}, \quad D_m = (k_0 k_1 \dots k_m)^{-2}.$$

We wish to generalise these formulae.

2. Let $\alpha_1, \alpha_2, \dots, \alpha_l$ be real or complex constants chosen so that the polynomial

(6. 4) $$u(x) = (\alpha_1 - x)(\alpha_2 - x) \dots (\alpha_l - x)$$

is real and non-negative in $\langle\alpha, \beta\rangle$. This will be, for instance, the case when the α_i are sufficiently large positive. We assume, moreover, that the determinants

(6. 5) $$\Delta_m = \begin{vmatrix} p_m(\alpha_1) & p_{m+1}(\alpha_1) & \dots & p_{m+l-1}(\alpha_1) \\ p_m(\alpha_2) & p_{m+1}(\alpha_2) & \dots & p_{m+l-1}(\alpha_2) \\ \cdot & & \dots & \cdot \\ p_m(\alpha_l) & p_{m+1}(\alpha_l) & \dots & p_{m+l-1}(\alpha_l) \end{vmatrix}, \quad m = 0, 1, 2, \dots, N,$$

are positive. Then the orthonormal polynomials $q_m(x)$ associated with the weight function $u(x)\, w(x)$ over $\langle\alpha, \beta\rangle$ are, for $n = 0, 1, 2, \dots, N-1$, given by the formula

(6. 6) $$u(x) q_m(x) = \left(\frac{k_m}{k_{m+l} \Delta_m \Delta_{m+1}}\right)^{1/2} \begin{vmatrix} p_m(x) & p_{m+1}(x) & \dots & p_{m+l}(x) \\ p_m(\alpha_1) & p_{m+1}(\alpha_1) & \dots & p_{m+l}(\alpha_1) \\ \cdot & & \dots & \cdot \\ p_m(\alpha_l) & p_{m+1}(\alpha_l) & \dots & p_{m+l}(\alpha_l) \end{vmatrix}.$$

For the proof cf. **4**, Theorem 2. 5, pp. 28—29, where the orthogonality of these polynomials is shown. As for the normalisation we note that the highest term of $q_m(x)$ is $(k_m k_{m+l})^{1/2} (\Delta_m / \Delta_{m+1})^{1/2} x^m + \dots$, and that

(6. 7) $$u(x)\, q_m(x) = \left(\frac{k_m}{k_{m+l}}\right)^{1/2} \left(\frac{\Delta_{m+1}}{\Delta_m}\right)^{1/2} p_m(x) + A_1 p_{m+1}(x) + \dots + A_l p_{m+l}(x).$$

Hence

(6. 8) $$\int_\alpha^\beta q_m^2(x)\, u(x)\, w(x)\, dx = \int_\alpha^\beta \left(\frac{k_m}{k_{m+l}}\right)^{1/2} \left(\frac{\Delta_{m+1}}{\Delta_m}\right)^{1/2} p_m(x)\, (k_m k_{m+l})^{1/2} \left(\frac{\Delta_m}{\Delta_{m+1}}\right)^{1/2} x^m w(x)\, dx = 1.$$

---

[7]) Cf. **4**, (2. 2 9), p. 26.

3. Let $h(x)$ be a given polynomial with real coefficients. We want to determine the extrema of the quotient

$$(6.9) \qquad \int_\alpha^\beta h(x) f^2(x) w(x)\, dx : \int_\alpha^\beta f^2(x) w(x)\, dx,$$

where the coefficients of the polynomial $f(x)$, of degree m, take arbitrary real values $u_0, u_1, \ldots, u_m$ not all zero. These maxima and minima are then characterised as the greatest and least zeros of the discriminant $H_m(\varrho)$ of the quadratic form (in the u_i)

$$(6.10) \qquad \int_\alpha^\beta (h(x)-\varrho)\,[u_0+u_1x+\ldots+u_mx^m]^2\, w(x)\, dx.$$

In order to compute $H_m(\varrho)$ we choose first the real numbers ε and ϱ so that $u(x)=\varepsilon(h(x)-\varrho)$ satisfies the above conditions. The highest coefficient of $u(x)$ in (6.4) has to be $(-1)^l$, so that ε depends only on the highest coefficient of h. By (6.3),

$$(6.11) \qquad \varepsilon^{-(m+1)} H_m(\varrho) = (k_0' k_1' \ldots k_m')^{-2},$$

where k_i' is the highest coefficient of the orthonormal polynomial $q_i(t)$ associated with the weight function $u(x)\,w(x)$. By (6.6)

$$(6.12) \qquad k_i' = (k_i k_{i+l})^{1/2} (\varDelta_i/\varDelta_{i+1})^{1/2}$$

so that

$$(6.13) \qquad \varepsilon^{-(m+1)} H_m(\varrho) = (k_0 k_1 \ldots k_m)^{-1} (k_l k_{l+1} \ldots k_{l+m})^{-1} \varDelta_{m+1}/\varDelta_0.$$

The quotient $\varDelta_{m+1}/\varDelta_0$ is a symmetric polynomial in $\alpha_1, \alpha_2, \ldots, \alpha_l$ which are the roots of $h(x)-\varrho$. Hence it is a polynomial of degree $m+1$ in ϱ, and *the equation* (6.13) *is an identity in* ϱ. The greatest and least zeros of the polynomial $\varDelta_{m+1}/\varDelta_0$ in ϱ yield the extrema in question.

4. The two simplest cases are $l=1$ and $l=2$ (compare (6.5)). If $l=1$, we have

$$(6.14) \qquad \varDelta_{m+1}/\varDelta_0 = k_0^{-1} p_{m+1}(\alpha_1).$$

If $l=2$, then

$$(6.15) \qquad \frac{\varDelta_{m+1}}{\varDelta_0} = \frac{p_{m+1}(\alpha_1)\,p_{m+2}(\alpha_2) - p_{m+1}(\alpha_2)\,p_{m+2}(\alpha_1)}{k_0 k_1 (\alpha_2-\alpha_1)} = (k_0 k_1)^{-1} K_{m+1}(\alpha_1, \alpha_2),$$

where K_m is the 'kernel function' [cf. **4,** (3.2.3), p. 42].

§ 7. The coefficients b_4 and b_5.

1. In the case b_4 we have

$$b_4 : b_1 = \int_{-1}^{1} P(x)\, U_3(x)\, (1-x^2)^{1/2}\, dx : \int_{-1}^{1} P(x)\, (1-x^2)^{1/2}\, dx. \tag{7.1}$$

By (1.2), it suffices to consider $\overline{B}(4, n)$ in each of the cases

$$\begin{aligned} &P(x) = A^2,\ P(x) = (1-x^2)\, B^2 && (n \text{ odd};\\ &P(x) = (1 \pm x)\, C^2 && (n \text{ even}). \end{aligned} \tag{7.2}$$

Now $U_3(x) = 8x^3 - 4x$, so that, on using the method of § 6, we are in the case $l = 3$, and we have to solve the equation $\Delta_{m+1}/\Delta_0 = 0$, that is an equation of the form

$$\frac{1}{(\alpha_1 - \alpha_2)(\alpha_2 - \alpha_3)(\alpha_3 - \alpha_1)} \begin{vmatrix} p_{m+1}(\alpha_1) & p_{m+2}(\alpha_1) & p_{m+3}(\alpha_1) \\ p_{m+1}(\alpha_2) & p_{m+2}(\alpha_2) & p_{m+3}(\alpha_2) \\ p_{m+1}(\alpha_3) & p_{m+2}(\alpha_3) & p_{m+3}(\alpha_3) \end{vmatrix} = 0, \tag{7.3}$$

where the $\alpha_1, \alpha_2, \alpha_3$ are the roots of $U_3(x) - \varrho = 0$; $m = \frac{1}{2}(n-1), \frac{1}{2}(n-3), \frac{1}{2}(n-2)$, respectively; and the polynomials $p_k(x)$ are associated with the weights $(1-x^2)^{1/2}$, $(1-x^2)^{3/2}$, $(1 \pm x)(1-x^2)^{1/2}$, respectively.

If we denote the maximal ϱ by $U_3(\zeta)$ where $-1 \leqq \zeta \leqq 1$, then

$$U_3(x) - \varrho = (x - \zeta)(8x^2 + 8x\zeta + 8\zeta^2 - 4), \tag{7.4}$$

so that $\alpha_1, \alpha_2, \alpha_3$ have the values

$$\zeta,\ \frac{1}{2}\left(-\zeta \pm \sqrt{2 - 3\zeta^2}\right). \tag{7.5}$$

Inserting these values in (7.3) we obtain an equation in ζ.

2. The case b_5 is in some respect even simpler. Here we may assume n odd and $P(x) = Q(x^2)$ where $\widehat{Q} = n' = \frac{1}{2}(n-1)$ (compare § 5).

Since $U_4(x) = 16x^4 - 12x^2 + 1$ we have

$$b_4 : b_1 = \int_0^1 Q(t)\,(16t^2 - 12t + 1)(1-t)^{1/2} t^{-1/2}\, dt : \int_0^1 Q(t)\,(1-t)^{1/2} t^{-1/2}\, dt. \tag{7.6}$$

Now putting $s = 2t - 1$, $Q(t)$ becomes a polynomial $Q^*(s)$ non-negative in $\langle -1, 1 \rangle$. Applying the theorem of Lukács, we find that we may restrict $Q(t)$ to the subclasses

$$\begin{aligned} &A^2(t), \quad t(1-t)\, B^2(t) && (n' \text{ even});\\ &tC^2(t), \quad (1-t)\, D^2(t) && (n' \text{ odd}). \end{aligned} \tag{7.7}$$

We are in the case $l = 2$, and the equations to be solved are, by (6.15) of the form

$$\frac{1}{\alpha_1 - \alpha_2} \begin{vmatrix} p_{m+1}(\alpha_1) & p_{m+2}(\alpha_1) \\ p_{m+1}(\alpha_2) & p_{m+2}(\alpha_2) \end{vmatrix} = 0, \tag{7.8}$$

where the $p_k(t)$ are associated with the weights $(1-t)^{1/2}t^{-1/2}$, $(1-t)^{3/2}t^{1/2}$, $(1-t)^{1/2}t^{1/2}$, $(1-t)^{3/2}t^{-1/2}$, respectively; and m being $\frac{1}{2}n'$, $\frac{1}{2}(n'-2)$, $\frac{1}{2}(n'-1)$, respectively. Also α_1 and α_2 are the roots of $16t^2-12t+1-\varrho$. Hence, putting $\varrho=16\tau^2-12\tau+1$, these roots are τ and $\frac{3}{4}-\tau$, so that (7.8) becomes

$$\frac{1}{2\tau-\frac{3}{4}}\begin{vmatrix} p_{m+1}(\tau) & p_{m+2}(\tau) \\ p_{m+1}\left(\frac{3}{4}-\tau\right) & p_{m+2}\left(\frac{3}{4}-\tau\right) \end{vmatrix}=0, \tag{7.9}$$

or $K_{m+1}\left(\tau, \frac{3}{4}-\tau\right)=0$ [compare (6.15)].

References.

1. R. P. Boas Jr.—M. Kac, Inequalities for Fourier transforms of positive functions, *Duke Math. Journal*, **12** (1945), pp. 189-206.
2. L. Fejér, Über trigonometrische Polynome, *Journal für die reine und angewandte Math.*, **146** (1915), pp. 53—82.
3. L. Fejér, Über eine Aufgabe der Harnackschen Potentialtheorie, *Göttinger Nachrichten* (1928), pp. 1—9.
4. G. Szegő, *Orthogonal Polynomials* (New York, 1939).

(Received September 9, 1949.)

King's College, Newcastle upon Tyne.
Stanford University.

Neuer Beweis zweier klassischer Sätze über Diophantische Approximationen.

Von OSKAR PERRON in München.

I. Von A. HURWITZ stammt der

Satz 1. *Zu jeder irrationalen Zahl α gibt es unendlich viele rationale Brüche x/y, für welche die Approximation gilt:*

$$\left|\alpha - \frac{x}{y}\right| < \frac{1}{\sqrt{5}\,y^2}.$$

Von KORKINE, ZOLOTAREFF und MARKOFF stammt der

Satz 2. *Sind a, b, c reelle Zahlen und ist $b^2 - 4ac = 1$, so gibt es unendlich viele Paare ganzer rationaler Zahlen x, y, für welche die Ungleichung gilt:*

$$|ax^2 + bxy + cy^2| \leq \frac{1}{\sqrt{5}}.$$

Der Satz 1 ist nicht etwa als der Spezialfall $a = 0$, $b = -1$, $c = \alpha$ in Satz 2 enthalten. Denn in diesem Spezialfall hat die Ungleichung von Satz 2 die unendlich vielen trivialen Lösungen $x =$ beliebig, $y = 0$, während Satz 1 unendlich viele Lösungen mit $y \neq 0$ behauptet. Zu beachten ist auch, daß in Satz 1 das Zeichen $<$ steht, während in Satz 2 beispielsweise bei der Form

$$\frac{1}{\sqrt{5}}x^2 + \frac{1}{\sqrt{5}}xy - \frac{1}{\sqrt{5}}y^2$$

offenbar nur Gleichheit erreichbar ist.

II. Für beide Sätze gibt es heute mehrere Beweise, und zwar wird Satz 1 in **[1]**, **[2]**, **[5]**, **[6]** mit Hilfe von Kettenbrüchen bewiesen[1]), in [3], **[7]** ohne Kettenbrüche. Satz 2 wird in **[4]** mit und in **[8]** ohne Kettenbrüche bewiesen. Hier soll für beide Sätze noch je ein einfacher Beweis ohne Kettenbrüche gegeben werden. Dabei wird das folgende Lemma gebraucht:

---

1) Die fetten Zahlen in eckigen Klammern weisen auf das Literaturverzeichnis am Schluß der Arbeit.

Lemma. *Sind δ, λ zwei Zahlen zwischen* 0 *und* 1, *so ist von den drei offenbar positiven Zahlen*

$$\delta,\ \lambda+\delta\lambda^2,\ 1-\lambda-\delta(1-\lambda)^2$$

wenigstens eine kleiner als $\frac{1}{\sqrt{5}}$. *Nur in dem Sonderfall* $\delta=\frac{1}{\sqrt{5}}$, $\lambda=\frac{3-\sqrt{5}}{2}$, *also niemals für rationales* λ, *sind alle drei gleich* $\frac{1}{\sqrt{5}}$.

III. Beweis des Lemmas. Man nehme an, es sei

$$\delta \geqq \frac{1}{\sqrt{5}}, \tag{1}$$

$$\lambda+\delta\lambda^2 \geqq \frac{1}{\sqrt{5}}, \tag{2}$$

$$1-\lambda-\delta(1-\lambda)^2 \geqq \frac{1}{\sqrt{5}}. \tag{3}$$

Addiert man dann die mit $(1-\lambda)^2$ multiplizierte Ungleichung (1) zu (3), so kommt $1-\lambda \geqq [(1-\lambda)^2+1]/\sqrt{5}$ und nach leichter Umformung:

$$\left(1-\lambda-\frac{\sqrt{5}}{2}\right)^2 \leqq \frac{1}{4}.$$

Daher ist $\lambda+\frac{1}{2}\sqrt{5}-1 \leqq \frac{1}{2}$, oder also

$$\lambda \leqq \frac{3-\sqrt{5}}{2}. \tag{4}$$

Multipliziert man (2) mit $(1-\lambda)^2$ und (3) mit λ^2, so folgt durch Addition $\lambda(1-\lambda) \geqq [(1-\lambda)^2+\lambda^2]/\sqrt{5}$, oder nach leichter Umformung:

$$\left(1-\lambda-\frac{\sqrt{5}}{2}\lambda\right)^2 \leqq \frac{1}{4}\lambda^2.$$

Daher ist $1-\lambda-\frac{\sqrt{5}}{2}\lambda \leqq \frac{1}{2}\lambda$, oder also

$$\lambda \geqq \frac{2}{3+\sqrt{5}} = \frac{3-\sqrt{5}}{2}. \tag{5}$$

Aus (4) und (5) zusammen folgt

$$\lambda = \frac{1}{2}(3-\sqrt{5}),\ \text{also}\ 1-\lambda = \frac{1}{2}(\sqrt{5}-1), \tag{6}$$

und nach (3) ist dann

$$\frac{\sqrt{5}-1}{2} - \delta\left(\frac{\sqrt{5}-1}{2}\right)^2 \geqq \frac{1}{\sqrt{5}},$$

oder also

$$\delta\left(\frac{\sqrt{5}-1}{2}\right)^2 \leqq \frac{\sqrt{5}-1}{2} - \frac{1}{\sqrt{5}} = \frac{3-\sqrt{5}}{2\sqrt{5}} = \frac{1}{\sqrt{5}}\left(\frac{\sqrt{5}-1}{2}\right)^2.$$

Zusammen mit (1) ergibt das:

(7) $$\delta = \frac{1}{\sqrt{5}}.$$

Wenn also nicht (6) und (7) gilt, muß wenigstens eine der drei Annahmen (1), (2), (3) falsch sein. Damit ist das Lemma bewiesen.

IV. Beweis von Satz I. Mit Hilfe des Dirichletschen Schubladenprinzips schließt man in bekannter Weise sofort auf die Existenz unendlich vieler Paare ganzer Zahlen x, y, für die $\left|\alpha - \frac{x}{y}\right| < \frac{1}{y^2}$ ist. Triviälerweise kann man x, y teilerfremd und $y > 0$ annehmen. Setzt man demgemäß

(8) $$\alpha - \frac{x}{y} = \varepsilon \frac{\delta}{y^2},$$

wo $0 < \delta < 1$ und $\varepsilon = \pm 1$ ist, und bestimmt zwei ganze Zahlen x_1, y_1 (eindeutig) so, daß

(9) $$x y_1 - y x_1 = \varepsilon, \qquad 0 < y_1 < y$$

ist, so hat man zunächst

$$\left|\alpha - \frac{x_1}{y_1}\right| = \left|\alpha - \frac{x}{y} + \frac{x}{y} - \frac{x_1}{y_1}\right| \leqq \left|\alpha - \frac{x}{y}\right| + \frac{1}{y y_1} < \frac{1}{y^2} + \frac{1}{y},$$

$$\left|\alpha - \frac{x - x_1}{y - y_1}\right| = \left|\alpha - \frac{x}{y} + \frac{x}{y} - \frac{x - x_1}{y - y_1}\right| \leqq \left|\alpha - \frac{x}{y}\right| + \frac{1}{y(y - y_1)} < \frac{1}{y^2} + \frac{1}{y}.$$

Die linken Seiten sind also beliebig klein, wenn nur y groß genug ist, und daraus folgt, daß es entsprechend den unendlich vielen verschiedenen Brüchen $\frac{x}{y}$ auch unter den Brüchen $\frac{x_1}{y_1}$ unendlich viele verschiedene gibt, und ebenso unter den Brüchen $\frac{x - x_1}{y - y_1}$.

Definiert man weiter zwei Zahlen δ_1, δ_2 durch

(10) $$\alpha - \frac{x_1}{y_1} = \varepsilon \frac{\delta_1}{y_1^2}, \quad \alpha - \frac{x - x_1}{y - y_1} = -\varepsilon \frac{\delta_2}{(y - y_1)^2},$$

so ist mit Rücksicht auf (8) und (9)

$$\varepsilon \delta_1 = y_1^2 \left(\alpha - \frac{x_1}{y_1}\right) = y_1^2 \left(\frac{\varepsilon \delta}{y^2} + \frac{x}{y} - \frac{x_1}{y_1}\right) = y_1^2 \left(\frac{\varepsilon \delta}{y^2} + \frac{\varepsilon}{y y_1}\right),$$

also, wenn $\frac{y_1}{y} = \lambda$ gesetzt wird:

$$\delta_1 = \lambda + \delta \lambda^2.$$

Ferner ist, wieder mit Rücksicht auf (8) und (9)

$$-\varepsilon \delta_2 = (y - y_1)^2 \left(\alpha - \frac{x - x_1}{y - y_1}\right) = (y - y_1)^2 \left(\frac{\varepsilon \delta}{y^2} + \frac{x}{y} - \frac{x - x_1}{y - y_1}\right) =$$

$$= (y - y_1)^2 \left(\frac{\varepsilon \delta}{y^2} - \frac{\varepsilon}{y(y - y_1)}\right),$$

also, da $\frac{y-y_1}{y} = 1-\lambda$ ist:

$$\delta_2 = 1-\lambda-\delta(1-\lambda)^2.$$

Aus (8) und (10) hat man also:

$$\left|\alpha-\frac{x}{y}\right| = \frac{\delta}{y^2}, \quad \left|\alpha-\frac{x_1}{y_1}\right| = \frac{\lambda+\delta\lambda^2}{y_1^2}, \quad \left|\alpha-\frac{x-x_1}{y-y_1}\right| = \frac{1-\lambda-\delta(1-\lambda)^2}{(y-y_1)^2}.$$

Da aber $\lambda = \frac{y_1}{y}$ rational ist und ebenso wie δ zwischen 0 und 1 liegt, ist nach dem Lemma von den drei rechts auftretenden Zählern wenigstens einer kleiner als $\frac{1}{\sqrt{5}}$.

V. Beweis von Satz 2. Für $a=0$, also $b=\pm 1$ ist nichts zu beweisen, weil die Ungleichung des Satzes dann die unendlich vielen Lösungen $x=$ beliebig, $y=0$ hat[2]). Sei daher jetzt $a \neq 0$. Wegen $b^2-4ac=1$ ist dann

$$\begin{aligned} ax^2+bxy+cy^2 &= a\left(x+\frac{b+1}{2a}y\right)\left(x+\frac{b-1}{2a}y\right) = \\ &= -a\left(\frac{1+b}{2a}y+x\right)\left(\frac{1-b}{2a}y-x\right). \end{aligned} \tag{11}$$

Ist $\frac{1-b}{2a}$ rational, so kann man die letzte Klammer durch unendlich viele (zueinander proportionale) Paare x, y zu null machen, und es ist nichts mehr zu beweisen. Daher sei jetzt $\frac{1-b}{2a}$ irrational. Setzt man dann

$$\frac{1-b}{2a}y-x = \frac{\varepsilon\delta}{y}, \quad \text{also} \quad 2ax+by = y-\frac{2a\varepsilon\delta}{y}, \tag{12}$$

wo $y>0$, $\delta>0$, $\varepsilon=\pm 1$ ist, so gibt es nach Satz 1 unendlich viele Paare x, y, für die $\delta < 1/\sqrt{5}$ ist.

Führt man jetzt wieder wie beim Beweis von Satz 1 die Zahlen x_1, y_1 ein und setzt

$$ax^2+bxy+cy^2 = -\varepsilon\delta', \tag{13}$$

$$ax_1^2+bx_1y_1+cy_1^2 = -\varepsilon\delta_1', \tag{14}$$

$$a(x-x_1)^2+b(x-x_1)(y-y_1)+c(y-y_1)^2 = \varepsilon\delta_2', \tag{15}$$

so wird Satz 2 bewiesen sein, wenn sich zeigen läßt, daß bei genügend großem y von den drei Zahlen $|\delta'|$, $|\delta_1'|$, $|\delta_2'|$ wenigstens eine $\leqq \frac{1}{\sqrt{5}}$ ist.

---

[2]) Übrigens hat sie auch unendlich viele Lösungen mit $y \neq 0$, was für rationales α trivial ist und für irrationales α aus Satz 1 folgt.

Nun folgt zunächst aus (13), (11) und (12):

$$\varepsilon\delta' = a\left(\frac{1+b}{2a}y + \frac{1-b}{2a}y - \frac{\varepsilon\delta}{y}\right)\frac{\varepsilon\delta}{y} = \varepsilon\delta - \frac{a\varepsilon^2\delta^2}{y^2};$$

daher:

$$\delta' = \delta - \frac{a\varepsilon\delta^2}{y^2}.$$

Somit liegt für genügend großes y der Quotient $\frac{\delta'}{\delta}$ beliebig nahe bei 1, so daß wegen $0 < \delta < \frac{1}{\sqrt{5}}$ insbesondere $0 < \delta' < 1$ ist.

Aus (13) und (14) folgt durch Elimination von c unter Berücksichtigung von (9)

$$\varepsilon(\delta_1' y^2 - \delta' y_1^2) = a(x^2 y_1^2 - x_1^2 y^2) + b(xyy_1^2 - x_1 y_1 y^2) =$$
$$= a\varepsilon(xy_1 + x_1 y) + b\varepsilon y y_1 = a\varepsilon(2xy_1 - \varepsilon) + b\varepsilon y y_1 = (2ax + by)\varepsilon y_1 - a.$$

Daher

$$\delta_1' = \delta'\left(\frac{y_1}{y}\right)^2 + \frac{2ax+by}{y}\cdot\frac{y_1}{y} - \frac{a\varepsilon}{y^2} = \delta'\left(\frac{y_1}{y} + \frac{2ax+by}{2\delta' y}\right)^2 - \frac{(2ax+by)^2}{4\delta' y^2} - \frac{a\varepsilon}{y^2}.$$

Mit Rücksicht auf (13) ist aber

$$(2ax+by)^2 = 4a(ax^2 + bxy) + b^2y^2 = 4a(-cy^2 - \varepsilon\delta') + b^2y^2 = y^2 - 4a\varepsilon\delta',$$

so daß die vorige Formel übergeht in:

$$\delta_1' = \delta'\left(\frac{y_1}{y} + \frac{2ax+by}{2\delta' y}\right)^2 - \frac{1}{4\delta'}.$$

Setzt man hier für $2ax+by$ den Ausdruck aus (12), so kommt schließlich:

$$\delta_1' = \delta'\left(\frac{y_1}{y} + \frac{1}{2\delta'} - \frac{a\varepsilon\delta}{\delta' y^2}\right)^2 - \frac{1}{4\delta'} = \delta'\left(\frac{y_1}{y} - \frac{a\varepsilon\delta}{\delta' y^2}\right)^2 + \left(\frac{y_1}{y} - \frac{a\varepsilon\delta}{\delta' y^2}\right).$$

Wenn man c aus (13) und (15) eliminiert, so führt eine ganz analoge Rechnung, wobei lediglich x_1, y_1, δ_1' durch $x_1 - x, y_1 - y, -\delta_2'$ zu ersetzen sind, zu der Formel

$$-\delta_2' = \delta'\left(\frac{y_1-y}{y} - \frac{a\varepsilon\delta}{\delta' y^2}\right)^2 + \left(\frac{y_1-y}{y} - \frac{a\varepsilon\delta}{\delta' y^2}\right).$$

Setzt man nun

(16) $$\frac{y_1}{y} - \frac{a\varepsilon\delta}{\delta' y^2} = \lambda,$$

so gehen die beiden letzten Formeln über in

(17) $$\delta_1' = \lambda + \delta'\lambda^2, \quad \delta_2' = 1 - \lambda - \delta'(1-\lambda)^2.$$

Nun liegt für genügend großes y, wie oben bemerkt, δ/δ' beliebig nahe bei 1, also λ nach (16) beliebig nahe bei y_1/y. Sollte also $\lambda \leqq 0$ sein, so ist $|\lambda|$ sehr klein, also nach (17) gewiß $|\delta_1'| < 1/\sqrt{5}$. Sollte aber $\lambda \geqq 1$ sein, so ist $|\lambda - 1|$ sehr klein, nach (17) also $|\delta_2'| < 1/\sqrt{5}$. Daher bleibt nur noch der Fall $0 < \lambda < 1$ zu untersuchen; in diesem ist aber nach dem Lemma von den drei (positiven) Zahlen δ', δ_1', δ_2' wenigstens eine $\leqq 1/\sqrt{5}$. Damit ist nun alles bewiesen.

Literatur.

[1] E. Borel, Contribution à l'analyse arithmétique du continu, *Journal de math. pures et appliquées,* (5) **9** (1903), S. 329 - 375.

[2] A. Hurwitz, Über die angenäherte Darstellung der Irrationalzahlen durch rationale Brüche, *Math. Annalen,* **39** (1891), S. 279—284, oder Werke, Bd. II, S. 122 – 128.

[3] A. Khintchine, Neuer Beweis und Verallgemeinerung eines Hurwitzschen Satzes, *Math. Annalen,* **111** (1935), S. 631—637.

[4] A. A. Markoff, Sur les formes quadratiques binaires indéfinies, *Math. Annalen,* **15** (1879), S. 381—406.

[5] O. Perron, *Die Lehre von den Kettenbrüchen* (Leipzig, 1913; zweite Auflage 1929), § 14.

[6] O. Perron, *Irrationalzahlen* (Leipzig, 1921; dritte Auflage 1947), § 36.

[7] O. Perron, Über einen Approximationssatz von Hurwitz und über die Approximation einer komplexen Zahl durch Zahlen des Körpers der dritten Einheitswurzeln, *Sitzungsberichte der Bayerischen Akademie der Wissenschaften, math.-naturwiss. Abteilung,* 1931, S. 129—154.

[8] O. Perron, Eine Abschätzung für die untere Grenze der absoluten Beträge der durch eine reelle oder imaginäre binäre quadratische Form darstellbaren Zahlen, *Math. Zeitschrift,* **35** (1932), S. 563—578.

(Eingegangen am 1. September, 1949.)

Sur les zéros des polynomes associés à un polynome.

Par PAUL MONTEL à Paris.

1. Etant donnés deux polynomes à coefficients réels $P(z)$ et $Q(z)$, on peut leur faire correspondre le polynome à coefficients complexes

$$\pi(z) = P(z) + iQ(z).$$

Inversement, étant donné un polynome $\pi(z)$ à coefficients complexes, on peut le représenter d'une manière unique sous la forme précédente. Nous dirons que les polynomes $P(z)$ et $Q(z)$ sont les polynomes associés au polynome $\pi(z)$ et nous appellerons faisceau associé à $\pi(z)$, le faisceau des polynomes $P(z) + \lambda Q(z)$, λ désignant un nombre réel.

Nous nous proposons d'examiner la nature des liens qui unissent les zéros de $\pi(z)$ à ceux de ses associés. Certains sont déjà connus; par exemple, un théorème de M. FUJIWARA montre que si tous les zéros de $\pi(z)$ sont situés dans le demi-plan $\Im z > 0$, ou dans le demi-plan $\Im z < 0$, les zéros de $P(z)$ et de $Q(z)$ sont réels, distincts et entrelacés; et réciproquement[1]).

2. Nous dirons qu'un zéro est supérieur ou inférieur suivant que sa partie imaginaire est positive ou négative. Supposons que le polynome $\pi(z)$, de degré n, admette p zéros supérieurs et $q = n - p$ zéros inférieurs. Nous pouvons toujours supposer $p \geqq q$ en remplaçant au besoin $Q(z)$ par $-Q(z)$. Soit ζ, un point du demi-plan supérieur $\Im z > 0$. Je dis que le nombre des zéros supérieurs de $P(z) + \zeta Q(z)$ est invariable et donc toujours égal à p. En effet, les zéros de ce polynome varient d'une manière continue avec ζ; l'un d'eux ne peut passer du plan supérieur au plan inférieur qu'en traversant l'axe des quantités réelles. Il y aurait donc un nombre réel x qui serait racine d'une équation de la forme

$$P(z) + \zeta Q(z) = 0,$$

ζ désignant un nombre complexe, ce qui est impossible. Donc, lorsque le point ζ se déplace dans le demi-plan supérieur en décrivant une courbe

---

[1]) M. FUJIWARA, Einige Bemerkungen über die elementare Theorie der algebraischen Gleichungen, *Tôhoku Math. Journal,* **9** (1916), p. 102—108.

continue qui ne rencontre pas l'axe réel, le nombre des zéros supérieurs de $P(z)+\zeta Q(z)$ est constant et égal à p.

Supposons maintenant que le point ζ décrive le segment de droite qui unit le point i à un point de l'axe réel d'abscisse λ. Tous les zéros varient d'une manière continue avec ζ et, lorsque ζ est venu en λ, les zéros sont devenus réels ou imaginaires conjugués. Comme le nombre des zéros inférieurs est toujours égal à q, il ne peut y avoir que q couples de zéros imaginaires conjugués au plus: il y a donc au moins $p-q$ zéros réels. Je dis que ces points sont distincts. En effet, s'il n'y avait pas $p-q$ zéros réels et distincts, il y aurait $p-q-2$ points réels distincts au plus, un point réel zéro double et q couples de zéros imaginaires conjugués. Faisons varier λ infiniment peu; le zéro double donne deux zéros simples, nécessairement réels quisqu'il ne peut y avoir plus de q couples de zéros imaginaires conjugués. Mais cela est impossible car, le zéro double correspond à un point simple de la courbe $P(x)+yQ(x)=0$, puisque cette courbe, supposée indécomposable, n'a aucun point multiple à distance finie. Or, en un point simple, les zéros voisins ne peuvent rester réels pour toutes les valeurs voisines de y. Nous pouvons donc énoncer le théorème suivant:

Si un polynome admet p zéros supérieurs et q zéros inférieurs $(p>q)$ chacun des polynomes du faisceau associé admet au moins $p-q$ zéros réels et distincts.

Le nombre $p-q=r$ sera appelé l'indice du faisceau ou de le fraction rationnelle $P/Q=R$.

Pour $q=0$, on voit que $P(z)+\lambda Q(z)$ a tous ses zéros réels et distincts quel que soit λ, d'où l'on déduit aussitôt que les zéros de deux de ces polynomes sont entrelacés: c'est le théorème de M. FUJIWARA[2]).

3. Comment distinguer si un zéro réel de $P(z)+\lambda Q(z)$ vient du demi-plan supérieur ou du demi-plan inférieur lorsque le point ζ, situé dans le demi-plan supérieur, vient se placer en λ?

On peut, sans modifier les raisonnements et les résultats, remplacer $P(z)$ et $Q(z)$ par deux autres polynomes du faisceau: prenons par exemple

$$P_1(z)=P(z)+\lambda Q(z),\quad Q_1(z)=P(z)+\mu Q(z),$$

avec $\lambda>\mu$. On voit que

$$\pi_1(z)=P(z)+iQ(z)=(1+i)\left[P(z)+\frac{\lambda+i\mu}{1+i}Q(z)\right]=(1+i)\,[P(z)+\zeta Q(z)],$$

ζ étant représenté par un point du demi-plan supérieur. Donc $\pi(z)$ et $\pi_1(z)$ ont les mêmes propriétés.

Soit donc α, un zéro réel de $P(z)+\lambda Q(z)$; nous pouvons supposer

---

[2]) Voir aussi P. MONTEL, Sur les fractions rationnelles à termes entrelacés, *Mathematica*, **5** (1931), p. 110–129.

que c'est un zéro de $Q(z)$ en changeant au besoin de polynomes de base; nous pouvons aussi supposer que $\alpha=0$, en remplaçant au besoin x par $x+\alpha$. On a

$$P(z)=P(0)+zP'(0)+\ldots,\quad Q(z)=zQ'(0)+\ldots,$$

avec $P(0)\,Q'(0)\neq 0$. Lorsque $\frac{1}{\zeta}$ est voisin de 0, l'équation

$$\frac{1}{\zeta}P(z)+Q(z)=\frac{P(0)}{\zeta}+zQ'(0)+\ldots=0$$

a une racine voisine de 0, dont la valeur principale est

$$-\frac{1}{\zeta}\frac{P(0)}{Q'(0)}=-\frac{A}{\zeta},$$

A désignant le résidu de la fraction $\frac{P}{Q}$ pour le pôle 0. Donc: un zéro simple réel vient du demi-plan supérieur ou du demi-plan inférieur suivant que le résidu correspondant est positif ou négatif.

Si 0 est un zéro multiple, d'ordre m, on aura $Q(z)=z^m\frac{Q^{(m)}(0)}{m!}+\ldots$. L'équation s'écrit

$$\frac{P(0)}{\zeta}+z^m\frac{Q^{(m)}(0)}{m!}+\ldots=0;$$

on voit qu'il existe m racines voisines de zéro et qu'il y a toujours des zéros supérieurs et des zéros inférieurs. C'est pourquoi les r zéros réels sont distincts et, lorsque $q=0$, tous les zéros sont distincts.

4. Supposons maintenant que, le point ζ venant du demi-plan supérieur en un point λ de l'axe réel, le polynome $P(z)+\lambda Q(z)$ admette $\rho>r$ zéros réels. Les $\rho-r$ zéros réels autres que les r provenant de zéros supérieurs de $P(z)+\zeta Q(z)$, proviennent de $\frac{1}{2}(\rho-r)$ zéros supérieurs et de $\frac{1}{2}(\rho-r)$ zéros inférieurs qui sont venus se placer sur l'axe réel, car il doit toujours rester un nombre égal de zéros imaginaires dans les deux demi-plans puisque les $n-\rho$ zéros non réels sont deux à deux conjugués.

Voyons comment se présente dans les différents cas la décomposition en éléments simples de la fraction $R(z)$.

Si $q=0$, les n zéros réels et distincts de $Q(z)=0$ proviennent de zéros supérieurs, on a donc

$$R=A_0+\sum_{i=1}^{n}\frac{A_i}{z-\alpha_i},\qquad (\alpha_1<\alpha_2<\ldots<\alpha_n).$$

Tous les A_i $(i=1,2,\ldots,n)$ sont positifs. D'ailleurs, lorsque λ croît en restant réel, chaque zéro de $P+\lambda Q$ varie dans le même sens, sinon un même zéro correspondrait à deux valeurs différentes de λ, ce qui est impossible. Comme, dans le voisinage de α_i, $z-\alpha_i$ est comparable à $-A_i/\lambda$, si on fait croître λ à partir de $-\infty$, $z-\alpha_i$ est positif et z croît de α_i à α_{i+1} lorsque λ varie de

$-\infty$ à $+\infty$. Donc, chaque polynome a un zéro et un seul dans l'intervalle (α_i, α_{i+1}); les deux segments $(-\infty, \alpha_1)$ et $(\alpha_n, +\infty)$ sont considérés comme formant un intervalle unique. Pour deux valeurs de λ, les zéros des deux polynomes du faisceau sont entrelacés.

Réciproquement, supposons que les zéros de $P+\lambda Q$ soient toujours réels et distincts. Le développement de $R(z)$ en élements simples a la forme précédente et tous les A_i ont le même signe. Si, en effet, deux résidus A_k et A_{k+1} étaient de signes différents, $R'(z)$ s'annulerait pour une valeur β comprise entre α_k et α_{k+1}; on aurait

$$\frac{P(\beta)}{Q(\beta)} = \frac{P'(\beta)}{Q'(\beta)} = -\lambda$$

et β serait un zéro double de $P(z)+\lambda Q(z)$, ce qui est contre l'hypothèse. Donc P et Q ont leurs zéros réels, distincts et entrelacés. Ce résultat est dû à M. Kakeya[3]).

Dans le cas où le polynome $\pi(z)$ ne dépend que de z^s, on peut remplacer le demi-plan par un secteur limité par deux demi-droites faisant entre elles un angle de π/s. On obtient des résultats semblables aux précédents, les rayons d'une étoile de sommet origine et d'ouverture π/s se substituant à l'axe réel.

5. Examinons encore le cas où $q=1$, $p=n-1$. Si le point ζ décrit le segment qui joint le point i au point λ de l'axe réel, deux racines au plus restent imaginaires conjuguées. Il y a donc, pour $P(z)+\lambda Q(z)$, $n-2$ zéros réels et distincts et n zéros réels au plus. Dans le cas où il n'y a, quel que soit λ, que $n-2$ zéros réels, $R(z)$ s'écrit

$$R(z) = A_0 + \sum_{i=1}^{n-2} \frac{A_i}{z-\alpha_i} + I(z).$$

Si, pour une valeur de λ, il y a n zéros réels, on peut les supposer distincts en faisant au besoin varier légèrement λ; on peut supposer aussi, en faisant au besoin un changement de paramètre, que cela arrive pour $Q(z)$. Alors $R(z)$ prend la forme

$$R(z) = A_0 + \sum_{i=1}^{n-1} \frac{A_i}{z-\alpha_i} - \frac{A'_{n-1}}{z-\alpha'_{n-1}}.$$

Dans le premier cas, $I(z)$ désigne l'élément simple correspondant aux zéros conjugués. Tous les A_i sont positifs puisque tous les α_i viennent du demi-plan supérieur.

Dans le second cas, tous les A_i et A'_{n-1} sont positifs puisque α'_{n-1} seul vient du demi-plan inférieur. On verrait comme précédemment que $R'(z)$ s'annule pour au moins une valeur réelle et que, par conséquent, $R(z)+\lambda$

[3]) Cf. M. Fujiwara, loc. cit.

ne peut, quel que soit λ, avoir n zéros réels et distincts. D'ailleurs nous savons bien que, dans ce cas, $-A'_{n-1}$ doit être positif.

Les deux cas précédents peuvent se présenter. Le premier, par exemple, pour $\pi(z)=iz^3+z^2+(1+i)z+1$ et le second pour $\pi(z)=iz^3+z^2+(1-i)z-1$.

Dans le cas général, lorsque p et q sont fixés, nous pourrons avoir pour $R(z)$ les $q+1$ représentations suivantes. Nous supposons que l'on a pris pour $Q(z)$ l'un des polynomes du faisceau $P(z)+\lambda Q(z)$ qui a le plus de zéros réels et distincts. Nous désignons toujours par r la différence $p-q=n-2q$. Tous les numérateurs des fractions simples sont positifs; les expressions I_k désignent des fractions rationnelles admettant $2k$ pôles imaginaires conjugués et nulles à l'infini.

$$R(z)=A_0+\sum_{i=1}^{r}\frac{A_i}{z-\alpha_i}+I_q,$$

. . . .

$$R(z)=A_0+\sum_{i=1}^{r+h}\frac{A_i}{z-\alpha_i}-\sum_{j=r+1}^{r+h}\frac{A'_j}{z-\alpha'_j}+I_{q-h},$$

. . . .

$$R(z)=A_0+\sum_{i=1}^{r+q}\frac{A_i}{z-\alpha_i}-\sum_{j=r+1}^{r+q}\frac{A'_j}{z-\alpha'_j}.$$

Cette classification n'est pas purement théorique. On peut construire des exemples correspondant à chacune des représentations. En effet, les α_i et α'_j étant choisis arbitrairement, on peut les placer de manière que les résidus aient le signe que l'on veut; en remplaçant ensuite I_k par εI_k, on peut prendre ε assez petit pour que le fraction R se comporte comme pour I_k identiquement nul puisque cette fraction I_k a un module borné quel que soit z réel.

6. On remarquera que le seul cas où le nombre des zéros réels de $P(z)+\lambda Q(z)$ demeure invariable quelle que soit la valeur réelle de λ, est celui où le nombre de ces zéros est égal à r. On peut donc énoncer le théorème suivant:

Si le polynome $P(z)+\lambda Q(z)$ admet, quel que soit λ réel, $n-2k$ zéros réels, la fraction rationnelle $\frac{P}{Q}=R(z)$ est décomposable sous la forme

$$R(z)=A_0+\sum_{i=1}^{k}\frac{A_i}{z-\alpha_k}+I_k,$$

les A_i étant de même signe. Les zéros sont donc distincts et le polynome $\pi(z)=P+iQ$ a k zéros inférieurs et $n-k$ zéros supérieurs.

Pour $k=0$, on retrouve le théorème de M. Kakeya.

Si le nombre des zéros réels des polynomes du faisceau varie entre $n-2k$ et $n-2k'$ $(k'<k)$, la fraction R peut admettre plusieurs représentations

correspondant à des valeurs de q prises entre k et $\frac{n}{2}$. Si $n=2h$ ou $n=2h+1$, le nombre de ces représentations est égal à $h-k+1$.

7. Soient $P+iQ$, de degré n, admettant p zéros supérieurs et q zéros inférieurs et $\frac{P}{Q}=R$; P_1+iQ_1, de degré n_1, admettant p_1 zéros supérieurs et q_1 zéros inférieurs et $\frac{P_1}{Q_1}=S$. Formons la fraction composée $R[S(z)]$ qui est de degré nn_1. Posons $S(z)=\zeta$; l'équation

(1) $$R[S(z)]+\lambda+i\mu=0 \ (\mu>0) \quad \text{ou} \quad R(\zeta)+\lambda+i\mu=0$$

admet p zéros supérieurs $\zeta_1, \zeta_2, \ldots, \zeta_p$ et q zéros inférieurs $\zeta'_1, \zeta'_2, \ldots, \zeta'_q$. L'équation $S(z)-\zeta_i=0$ admet q_1 zéros supérieurs et p_1 inférieurs; au total, pq_1 supérieurs et pp_1 inférieurs. L'équation $S(z)-\zeta'_j=0$ admet p_1 zéros supérieurs et q_1 inférieurs; au total, p_1q supérieurs et q_1q inférieurs. Finalement, l'équation (1) a pp_1+qq_1 zéros inférieurs et pq_1+qp_1 supérieurs. Les nombres p' et q' relatifs à $R[S(z)]$ sont

$$p'=pq_1+qp_1, \quad q'=pp_1+qq_1, \quad \text{d'où} \quad p'-q'=(p-q)(q_1-p_1) \quad \text{et} \quad r'=rr_1,$$

Ainsi: *l'indice d'une fraction composée est le produit des indices des fractions composantes.* Le résultat est vrai quel que soit le nombre des fractions composantes. On peut remarquer que ce résultat est indépendant de l'ordre des compositions.

(Reçu le 17 septembre 1949)

Return to the self-adjoint transformation.

By E. R. LORCH in New York.

We return once more to the theory of the self-adjoint transformation in Hilbert space. This subject which was born shortly after the turn of the century, has since then attracted wide attention, — particularly so during the last twenty years. The central fact is the structure theorem which asserts the following:

If H is a self-adjoint transformation in Hilbert space $\mathfrak{H}$, then there exists a resolution of the identity $E(\lambda)$ such that the structure of H is completely summarized by the formula

$$H = \int_{-\infty}^{\infty} \lambda \, dE(\lambda). \tag{1}$$

Many proofs have been given of this theorem. Some are concerned with the bounded case. Others apply to the general situation. Still a third variety considers that the unbounded case is best treated by first carrying through a complete discussion for the bounded transformation. We shall not analyse the methods of these proofs, which are after all well known to those interested in this domain.

In the pages which follow we set down a new proof of this fundamental theorem. Our approach is to attack the most general (unbounded) situation directly from the start and to assume no knowledge of transformation theory except the most trivial facts. We should like to believe that our method yields the final result considerably more rapidly than those heretofore advanced. We obtain formula (1) in the following form:

Theorem. *Let H be a self-adjoint transformation in Hilbert space $\mathfrak{H}$. Let $\{\lambda_n\}$ be a set of real numbers $n=0, \pm 1, \pm 2, \ldots$, such that*

a) *for all n, $\lambda_n > \lambda_{n-1}$;* b) $\lim\limits_{n\to\infty} \lambda_n = \infty$; c) $\lim\limits_{n\to-\infty} \lambda_n = -\infty$.

Then there exists in $\mathfrak{H}$ a set of closed linear manifolds $\{\mathfrak{M}_n\}$, $n=0$, $\pm 1, \pm 2, \ldots$, orthogonal in pairs, spanning $\mathfrak{H}$, and such that H is defined on $\mathfrak{M}_n$ and satisfies

$$\lambda_n I \leqq H \leqq \lambda_{n+1} I. \tag{2}$$

It will be seen that this theorem gives formula (1) directly. In the first place the manifolds $\mathfrak{M}_\lambda$ define in an obvious way the projections $E(\lambda)$. Secondly, the inequality (2) coupled with elementary facts on orthogonality is precisely what makes it possible to define the integral of Riemann-Stieltjes type which is in (1). We remind the reader that $H \geqq \lambda I$ means $(Hf, f) \geqq \lambda (f, f)$.

Our methods are based on integrals around simple curves in the complex plane. The general form of these integrals is

$$K = \frac{1}{2\pi i} \int\limits_C \frac{f(\zeta)}{\zeta I - H} d\zeta. \tag{3}$$

These expressions are strongly reminiscent of the Cauchy integral formula, except that in the integrand we find instead of the usual $(\zeta - a)^{-1}$ the operator $(\zeta I - H)^{-1}$. It is easy to show that since H is self-adjoint the above operator is bounded providing ζ is not real. This means that our integrals are improper since the curve C cuts the real axis. The reason for this may be put another way. If the contour curve over which the integration is performed lies exclusively in the resolvent set of H we have an operator which was first considered briefly by F. Riesz and subsequently was used by the author to study the reducibility of normed rings, to investigate the theory of their radical, and to define an operational calculus of operators in general vector spaces. In the present work, the path of integration crosses the spectrum of H and hence the behavior of the integrand must be subject to a careful examination. The existence of our integrals rests on very special properties possessed by self-adjoint transformations. The fundamental functional equations satisfied by the integrals (3) are proved with the help of the Neumann expansion for the resolvent and, more important, the functional equation of the resolvent. All that is required to carry through the proofs is a little patience to determine that the operations on iterated integrals are legitimate. One further point may deserve mention: The ideas which underlie the procedure are quite transparent in spite on the fact that at times they may be slightly submerged under a technique and notation which are evoked by the present subject-matter as well as our approach to it. The reader with experience in this subject will wish to omit the reading of sections I and II.

I. Definitions.

Let T be a linear transformation which is defined for every vector f belonging to some linear set Δ which is dense in $\mathfrak{H}$. Consider the set of all pairs $\{g, g^*\}$ such that $(Tf, g) = (f, g^*)$ for all $f \in \Delta$. The relation $g \to g^*$ defines a linear transformation T^*, $T^* g = g^*$, called the adjoint of T. If the domain of T^* is identical with that of T and if $T = T^*$ on this common domain, we say that T is self-adjoint.

A transformation T is said to be closed if it has the following property: Let $\{f_n\}$ be a sequence of vectors belonging to the domain Δ of T. Suppose that $f_n \to f$ and $Tf_n \to g$. Then $f \in \Delta$ and $Tf = g$. It is very easy to see that every adjoint transformation T^* is closed. In particular, all self-adjoint transformations are closed.

II. The resolvent.

Let H be self-adjoint with a domain of definition Δ. Let ζ be a complex number, $\zeta = \alpha + \beta i$, with $\beta \neq 0$. Then if we write $(\zeta I - H)g = f$ for any arbitrary $g \in \Delta$ we have

$$|f|^2 = |(\alpha I - H)g|^2 + \beta^2 |g|^2,$$

hence $|g| \leqq |f|/|\beta|$. In particular, if $g \neq 0$, then $f \neq 0$, hence ζ is not in the point spectrum of H. If the totality of elements f obtained by allowing g to vary over Δ were not dense in H, then $\bar{\zeta} = \alpha - \beta i$ would be in the point spectrum of $H^* = H$. Finally, the elements f actually fill $\mathfrak{H}$. For if $\{g_n\}$ is a sequence such that $g_n \in \Delta$ and that $f_n \to f$, then $\{g_n\}$ is a convergent sequence since $|\beta| |g_n - g_m| \leq |f_n - f_m|$. Since H is a closed transformation, the element g to which $\{g_n\}$ converges belongs to Δ and $(\zeta I - H)g = f$. If we consolidate these facts we obtain the critically important result that *if ζ is not a real number, the transformation $(\zeta I - H)^{-1}$ exists and is a bounded linear transformation with a bound equal to or less than $|\beta|^{-1}$ where β is the imaginary part of ζ.*

If ζ is a fixed point in the complex plane which is not on the real axis and if ξ is a point near ζ, then we have

$$(4) \quad (\xi I - H)^{-1} = (\zeta I - H)^{-1} + (\zeta - \xi)(\zeta I - H)^{-2} + (\zeta - \xi)^2 (\zeta I - H)^{-3} + \ldots$$

This expansion is valid for all ξ such that $|\xi - \zeta| < 1/|(\zeta I - H)^{-1}|$. The relation (4) implies that the bound of $(\xi I - H)^{-1}$ is a continuous function of ξ for ξ lying off the real axis.

The functional equation for the resolvent, valid for non-real ζ and ξ, is

$$(5) \quad (\zeta I - H)^{-1} - (\xi I - H)^{-1} = (\xi - \zeta)(\zeta I - H)^{-1}(\xi I - H)^{-1}.$$

This may be established by multiplying both sides by $(\zeta I - H) = (\zeta - \xi) I + (\xi I - H)$.

III. The point spectrum.

If for a pair $\{\lambda, f\}$ with $f \neq 0$, we have $Hf = \lambda f$, then λ is said to be in the point spectrum of H. If the totality of such vectors f spans $\mathfrak{H}$, then H is said to have a pure point spectrum. Now due to the fact that characteristic vectors f_1 and f_2 which correspond to distinct characteristic values λ_1 and λ_2 are orthogonal to each other, the fundamental formula (1) is trivial for an H with pure point spectrum. In the methods which we employ in what follows the possible existence of the point spectrum causes annoyance.

For this reason, we shall remove it. If the space $\mathfrak{H}$ is separable, this is not necessary since the point spectrum contains at most a denumerable number of real numbers and in any neighborhood one may always cut the real axis at a point not a characteristic value of H. If $\mathfrak{H}$ is not separable, the procedure described below makes it unnecessary later always to amend our statements. We shall prove the following:

If H is a self-adjoint transformation in $\mathfrak{H}$, and if $\mathfrak{M}$ is the closed linear manifold spanned by the totality of characteristic vectors of H, then $\mathfrak{M}^\perp$ the orthogonal complement of $\mathfrak{M}$ reduces H, and in the space $\mathfrak{M}^\perp$, H is a self-adjoint transformation and has no point spectrum.

Proof: Let $\{\varphi_\alpha\}$ be an orthonormal set which spans $\mathfrak{M}$ and whose members are characteristic vectors of H. Let f be an arbitrary element in the domain $\varDelta$ of H. Suppose the expansion for f in terms of the set $\{\varphi_\alpha\}$ is $f \sim \Sigma a_\alpha \varphi_\alpha$. Let $f^* = \Sigma a_\alpha \varphi_\alpha$. Then since H is defined for f and for each φ_α it is defined for $f - \Sigma' a_\alpha \varphi_\alpha$ where Σ' denotes a sum containing only a finite number of terms. Since $H^2(\Sigma' a_\alpha \varphi_\alpha)$ is well defined and since

$$(H(f-\Sigma' a_\alpha \varphi_\alpha),\ H(\Sigma' a_\alpha \varphi_\alpha)) = (f-\Sigma' a_\alpha \varphi_\alpha,\ H^2(\Sigma' a_\alpha \varphi_\alpha)) = 0,$$

we have

$$|Hf|^2 = |H(f-\Sigma' a_\alpha \varphi_\alpha)|^2 + |H(\Sigma' a_\alpha \varphi_\alpha)|^2.$$

The sum $\Sigma a_\alpha \varphi_\alpha$ contains at most denumerably many non-zero terms. To simplify notation we shall write $\Sigma a_\alpha \varphi_\alpha = \lim\limits_{n\to\infty} \Sigma_1^n a_\alpha \varphi_\alpha$. The equation established immediately above shows that the sequence $\{H(\Sigma_1^n a_\alpha \varphi_\alpha)\}$ converges. Since H is a closed transformation, f^* and also $f - f^* = f - \Sigma a_\alpha \varphi_\alpha$ belong to the domain of H. Clearly $f - f^* \in \mathfrak{M}^\perp$.

Thus every element f in $\varDelta$ is the sum in a unique way of an element g in $\mathfrak{M}$ and h in $\mathfrak{M}^\perp$ where both g and h belong to $\varDelta$: $f = g + h$. Clearly, since $Hg \in \mathfrak{M}$ and since $(g, Hh) = (Hg, h) = 0$ we see that $Hh \in \mathfrak{M}^\perp$. Since the set $\varDelta$ of elements f is dense in $\mathfrak{H}$, the set of elements h is dense in $\mathfrak{M}^\perp$.

It remains to show that in $\mathfrak{M}^\perp$, H is self-adjoint. If there exists a pair $\{k, k^*\}$ in $\mathfrak{M}^\perp$ such that for every h in $\varDelta$ and in $\mathfrak{M}^\perp$, $(Hh, k) = (h, k^*)$, then for every f in $\varDelta$, $f = g + h$, $g \in \mathfrak{M}$, and $h \in \mathfrak{M}^\perp$, we have $(H(g+h), k) = (Hh, k) = (h, k^*) = (g+h, k^*)$. Since H is self-adjoint, $k \in \varDelta$ and $Hk = k^*$. This shows that considered in the space $\mathfrak{M}^\perp$, H is self adjoint. That H has no point spectrum in $\mathfrak{M}^\perp$ is obvious.

To establish (1) for the transformation H operating in $\mathfrak{H}$ it is sufficient to establish it for H in $\mathfrak{M}$ and for H in $\mathfrak{M}^\perp$. Since the case H in $\mathfrak{M}$ is trivial, there remains only the case H in $\mathfrak{M}^\perp$. In consequence, *from now on, by virtue of the preceding result, we shall assume that H has no point spectrum.*

IV. On a class of improper integrals.

We now introduce the integrals which are the key to the structure theorem. The path of integration is a curve C which lies in the complex plane. For the sake of convenience, C is assumed to be smooth except possibly for a few corners. The curve cuts the axis of reals at a non-zero angle at two point λ and μ. In addition, we shall assume that C is symmetric in the axis of reals. The letters m and n represent positive integers. The integral which concerns us is

$$K_{\lambda\mu}(m,n)=\frac{1}{2\pi i}\int_C(\zeta-\lambda)^m(\mu-\zeta)^n(\zeta I-H)^{-1}\,d\zeta. \tag{6}$$

The integral may be improper at $\zeta=\lambda$ or $\zeta=\mu$. We examine its behavior as follows. We alter C by eliminating from it two short segments which enclose these critical points. The resulting path of integration will be called D. Now the integral (6) over the path D is well defined and exists in the uniform topology of operators This is due to the fact that $(\zeta I-H)^{-1}$ is a continuous function of ζ (see equation (4) in section III). Now since $m\geq 1$ and since C cuts the axis of reals at a non-zero angle the operator $(\zeta-\lambda)^m(\zeta I-H)^{-1}$ is bounded near $\zeta=\lambda$. This type of argument proves that the integral (6) converges in the uniform topology and hence represents a bounded operator. Since the path of integration is symmetric about the real axis and due to the special structure of the integrand, it is clear that $K_{\lambda\mu}(m,n)$ is self-adjoint. Also by virtue of equation (4) in section II, it is clear that the value of the integral is not changed if the path is slightly deformed providing that the points λ and μ remain fixed. We now list these and the other properties of this operator which are of importance to us.

The improper integral $K_{\lambda\mu}(m,n)$ converges in the uniform topology and represents a bounded self-adjoint transformation. The value of the integral is not altered if the path of integration C is deformed slightly providing that the points $\zeta=\lambda$ and $\zeta=\mu$ remain on C.

The transformation $K_{\lambda\mu}(m,n)$ satisfies the further conditions:

a) $K_{\lambda\mu}(m,n)\cdot K_{\lambda\mu}(m',n')=K_{\lambda\mu}(m+m',n+n')$.

b) *If the intervals (λ,μ) and (λ',μ') have no points in common (or at most one end-point in common), then*

$$K_{\lambda\mu}(m,n)\cdot K_{\lambda'\mu'}(m',n')=0.$$

c) *The transformation $H-\lambda I$ is defined for every element in the range of $K_{\lambda\mu}(m,n)$ and furthermore*

$$(H-\lambda I)\cdot K_{\lambda\mu}(m,n)=K_{\lambda\mu}(m+1,n).$$

A similar statement may be made for the operator $\mu I-H$.

d) *For the elements in the range of* $K_{\lambda\mu}(m, n)$, H *is a bounded self-adjoint transformation which satisfies the inequalities* $\lambda I \leq H \leq \mu I$.

We start by proving a). We are concerned with the product of two integrals. We may assume that the associated paths of integration C and C' are such that one lies entirely within the other except at the two points λ and μ. Then using the functional equation for the resolvent (section II, equation (5)) we express the product of these integrals as an iterated integral. Making use of the Cauchy integral formula of ordinary function theory on the innermost integral we obtain a).

The proof of b) is similar to that of a). Since the two paths of integration lie each outside the other, the product of the two integrals yields zero.

To prove c) we note that the operator H is obviously defined for any element $(\zeta I - H)^{-1} f$. Since $H - \lambda I = (H - \zeta I) + (\zeta - \lambda) I$, the product of $K_{\lambda\mu}(m, n)$ by $H - \lambda I$ gives two integrals, one of which is $K_{\lambda\mu}(m+1, n)$ while the other is zero by the Cauchy theorem. Careful examination of this argument shows that use is made of the fact that H is a closed transformation. One may see this as follows. The integral $K_{\lambda\mu}(m, n)$ may be approximated by a finite Riemann sum which we designate by $\Sigma(m, n)$. If we multiply $\Sigma(m, n)$ by $H - \lambda I$ we obtain essentially an approximating sum $\Sigma(m+1, n)$ for $K_{\lambda\mu}(m+1, n)$. The result c) is obtained by taking limits in the sense of integration and using the closure of H.

To prove d), we note first that we may write the following relation for inner products in $\mathfrak{H}$; here f is an arbitrary vector in $\mathfrak{H}$:

$$\begin{aligned}(7)\quad ((H - \lambda I)\, K_{\lambda\mu}(m, n) f,\ K_{\lambda\mu}(m, n) f) &= (K_{\lambda\mu}(m+1, n) f,\ K_{\lambda\mu}(m, n) f) = \\ &= (K_{\lambda\mu}(m, n)\, K_{\lambda\mu}(m+1, n) f, f) = (K_{\lambda\mu}(2m+1, 2n) f, f).\end{aligned}$$

In writing this down, use has been made of relations a) and c). We shall show that there exists a self-adjoint transformation $L_{\lambda\mu}(2m+1, 2n)$ such that

$$(8)\qquad L^2_{\lambda\mu}(2m+1, 2n) = K_{\lambda\mu}(2m+1, 2n).$$

This will prove that the inner product in equation (7) is non-negative and hence that for $g = K_{\lambda\mu}(m, n) f$, $H \geq \lambda I$.

Consider to this effect the integral

$$(9)\qquad L_{\lambda\mu}(m, n) = \frac{1}{2\pi i} \int_C (\zeta - \lambda)^{m/2} (\mu - \zeta)^{n/2} (\zeta I - H)^{-1} d\zeta.$$

The square root of $(\zeta - \lambda)^m$ is understood to be that which is positive when $\zeta - \lambda$ is real and positive; that of $(\mu - \zeta)^n$ is positive when $\mu - \zeta$ is real and positive. By arguments similar to those used above, one may prove that the integral $L_{\lambda\mu}(m, n)$ exists and represents a bounded self-adjoint transformation; furthermore one may also establish equation (8). The integral (9) introduces a new feature in case $m = 1$; for in that case $(\zeta - \lambda)^{m/2} (\zeta I - H)^{-1}$

may have an infinite bound at $\zeta = \lambda$. A routine argument disposes of this difficulty.

This concludes the proof of the fact that $H \geq \lambda I$. The steps which demonstrate that $H \leq \mu I$ are similar.

V. On the linear manifolds associated with $K_{\lambda\mu}(m, n)$.

Associated with every transformation $K_{\lambda\mu}(m, n)$ we consider two closed linear manifolds: the manifold of the zeros of the transformation and the manifold of the closure of the range of the transformation. It is an obvious fact that each is the orthogonal complement of the other. In this section we develop some properties of these manifolds.

Let the closure of the range of $K_{\lambda\mu}(m, n)$ be denoted by $\mathfrak{M}_{\lambda\mu}(m, n)$. Let the set of its zeros be denoted by $\mathfrak{N}_{\lambda\mu}(m, n)$. We prove:

a) *$\mathfrak{M}_{\lambda\mu}(m, n)$ is independent of m and n; that is, $\mathfrak{M}_{\lambda\mu}(m, n) = \mathfrak{M}_{\lambda\mu}(m', n')$. This common manifold is denoted by $\mathfrak{M}_{\lambda\mu}$.*

b) *If λ, μ, ν are real numbers subject to $\lambda < \mu < \nu$ then $\mathfrak{M}_{\lambda\mu} + \mathfrak{M}_{\mu\nu} = \mathfrak{M}_{\lambda\nu}$ where the '$+$' operation is performed in the sense of addition of linear manifolds.*

To prove a), consider $\mathfrak{N}_{\lambda\mu}(m, n)$. We have $(H - \lambda I)\, K_{\lambda\mu}(m, n) = K_{\lambda\mu}(m+1, n)$. Since H has no point spectrum, this means that $\mathfrak{N}_{\lambda\mu}(m, n) = \mathfrak{N}_{\lambda\mu}(m+1, n)$. This type of reasoning yields a).

Toward b): By a) we may assume that all integers m, n, etc. are equal to 1. Also since the point spectrum is absent we may without changing the character of $\mathfrak{M}_{\lambda\mu}$ assume that it is defined by an integral operator whose integrand is $(2\pi i)^{-1}(\zeta - \lambda)(\zeta - \mu)(\zeta - \nu)(\zeta I - H)^{-1}$. The same is true of $\mathfrak{M}_{\mu\nu}$ and $\mathfrak{M}_{\lambda\nu}$. We denote the three operators by $T_{\lambda\mu}$, $T_{\mu\nu}$, and $T_{\lambda\nu}$. Now it is clear that a possible path of integration which defines $T_{\lambda\nu}$ is the sum of the paths defining $T_{\lambda\mu}$ and $T_{\mu\nu}$. Thus $T_{\lambda\nu} = T_{\lambda\mu} + T_{\mu\nu}$. Also, we have $T_{\lambda\mu} \cdot T_{\mu\nu} = 0$. These facts lead to the conclusion in b).

VI. Conclusion of proof.

We recall the method for proving our principal theorem. Let $\{\lambda_n\}$ be a monotone increasing set of real numbers such that $\lim_{n\to\infty} \lambda_n = \infty$ and $\lim_{n\to-\infty} \lambda_n = -\infty$. For every integer s and for some fixed m construct the transformation $K_{\lambda_s \lambda_{s+1}}(m, m)$. Determine the closure of its range $\mathfrak{M}_s$. The manifolds $\mathfrak{M}_s$ are orthogonal in pairs (in virtue of b) in section IV) and on $\mathfrak{M}_s$, $\lambda_s I \leq H \leq \lambda_{s+1} I$. To complete our proof it remains to show that the manifolds $\mathfrak{M}_s$ span $\mathfrak{H}$. This we proceed to do.

In the first place, the result b) in section V shows that instead of considering many manifolds, we need consider only one. Going beyond this, it is clear that we need consider only the single manifold $\mathfrak{P}_r$ which corresponds to the interval $(-r, r)$.

If the manifolds $\mathfrak{P}_r$, $0<r<\infty$, do not span $\mathfrak{H}$, there is a non-zero element f orthogonal to each $\mathfrak{P}_r$. We shall explicitly assume that f is orthogonal to each $\mathfrak{P}_r$ and subsequently prove that $f=0$. Now $\mathfrak{P}_r$ is the closure of the range of $K_{-r,r}(1,1)$. Since the latter operator is self-adjoint, we have $K_{-r,r}(1,1)f=0$ for all values of r. Thus we have for all r

$$(10) \qquad f=\frac{1}{2\pi i r^2}\int_C [(r^2-\zeta^2)\zeta^{-1}-(r^2-\zeta^2)(\zeta I-H)^{-1}]\,d\zeta\cdot f.$$

Here C is a circle with radius r and center the origin; the first term in the integrand yields f by Cauchy's theorem. Now a simple calculation shows that except for a constant factor the integrand in (10) is of the form $H(r^2-\zeta^2)\cdot\zeta^{-1}\cdot(\zeta I-H)^{-1}f$. If we remove the factor H from under the integral sign and perform the integration we obtain an element which we may denote by g_r. Hence equation (10) may be written in the form $f=Hg_r$. Now, by using (3) and standard techniques of evaluation of integrals, it is an easy matter to show that $|g_r|\leq 2r^{-1}|f|$. As $r\to\infty$, $g_r\to 0$ while $Hg_r\to f$. Since H is a closed transformation, $f=0$.

BARNARD COLLEGE, COLUMBIA UNIVERSITY.

(Received August 1, 1949.)

On limit periodic functions of infinitely many variables.

By HARALD BOHR in Copenhagen.

1. In the sequel the functions to be considered are continuous complex-valued functions of unrestricted real variables. Furthermore, convergence of a sequence of functions is always to be taken in the sense of uniform convergence over the whole range of the variable (or the variables).

2. Among the a.p. (almost periodic) functions of one variable x,

$$F(x) \sim \Sigma A_n e^{i\Lambda_n x},$$

the p.p. (purely periodic) functions $P(x)$ are the simplest ones; the periods of such a function are either all real numbers (in the case of $P(x)$ being constant), or the integral multiples of a real number $p_0 \neq 0$. Another simple, although more general case of a.p. functions $F(x)$ are the l.p. (limit periodic) functions $G(x)$ the set of which are obtained from the class of the p. p. functions $\{P(x)\}$ by closing it (with respect to uniform convergence), i. e.

$$\{G(x)\} = \mathrm{Cl}\,\{P(x)\}.$$

As easily seen (II, p. 141), for the l. p. functions $G(x)$ a kind of period still exists since two p.p. functions $P_1(x)$ and $P_2(x)$ which approximate a non-constant l.p. function $G(x)$ sufficiently well, must necessarily have periods with rational ratio. Denoting the class $\{G(x)\} = \mathrm{Cl}\,\{P(x)\}$ of all l. p. functions by C, and by $C_p\,(p \neq 0)$ the closure $\mathrm{Cl}\,\{P_p(x)\}$ of only those periodic functions $P_p(x)$ which have a rational multiple of p as one of their periods, we conclude that the set C is the union of all the sets C_p, i. e.

$$C = \bigcup_p C_p.$$

We may express this in the following way: We get the same set of functions (and not a smaller one) by closing first, for a fixed $p \neq 0$, the set $\{P_p(x)\}$ and then forming the union of all these closures, as we get by closing directly the whole set $\{P(x)\}$. As to the Fourier series of the l. p. functions, these are characterized, among the Fourier series of the a. p. functions, by having exponents Λ_n with mutually rational ratios; more particular, the exponents of a l. p. function from C_p are rational multiples of the number $\dfrac{2\pi}{p}$.

3. In the study of the a. p. functions of one variable the l. p. functions of an infinite number of variables play an important role (II, p. 118—163). We consider the enumerable-dimensional space of points $X=(x_1, x_2, \ldots)$ with arbitrary real coordinates where convergence of a sequence of points $X', X'', \ldots$ simply means convergence in each of the coordinates. A (continuous) function $P(X)=P(x_1, x_2, \ldots)$ is called p p. (purely periodic) with respect to the axis, if there exist non-vanishing real numbers $p_1, p_2, \ldots$ such that for each n the equation

$$P(x_1, x_2, \ldots, x_n+p_n, \ldots)=P(x_1, x_2, \ldots, x_n, \ldots)$$

holds good in the whole space. On account of the continuity of the function $P(X)$ we then also have (II, p. 135)

$$P(x_1+\nu_1 p_1, x_2+\nu_2 p_2, \ldots)=P(x_1, x_2, \ldots)$$

for each choice of the integers $\nu_1, \nu_2, \ldots$. By closing the set of all functions $P(X)$, p. p. with respect to the axis, we get the functions $G(X)$, l. p. with respect to the axis,

$$\{G(X)\}=\mathrm{Cl}\{P(X)\}.$$

Furthermore (II, p. 148), denoting for $p_1 \neq 0, p_2 \neq 0, \ldots$ by $C_{p_1, p_2, \ldots}$ the closure of only those of our p. p. functions $P_{p_1, p_2, \ldots}(X)$ which have rational multiples of $p_1, p_2, \ldots$ as periods with respect to the axis, we have just as before

$$C=\bigcup_{p_1, p_2, \ldots} C_{p_1, p_2, \ldots}.$$

Among the Fourier series of the a. p functions $F(X)$ of infinitely many variables

$$F(x_1, x_2, \ldots) \sim \Sigma A_n e^{i(\Lambda_{n,1}x_1+\Lambda_{n,2}x_2+\ldots+\Lambda_{n,m_n}x_{m_n})}$$

studied by Bochner (I), those belonging to l. p. functions of the class $C_{p_1, p_2, \ldots}$ are characterized by having, for each fixed m, all the numbers $\Lambda_{n,m}$ in the exponents equal to rational multiples of $\dfrac{2\pi}{p_m}$. Thus the Fourier series belonging to the l. p. functions of the class $C_{2\pi, 2\pi, \ldots}$ are just those Fourier series of a. p. functions for which the numbers $\Lambda_{n,m}$ are all rational.

4. We now introduce the notion of a substitution in our infinite-dimensional space as a linear one-to-one bicontinuous transformation T of the whole space on the whole space itself. As easily seen (III, p. 11 and V, p. 53) such a substitution may be written in the form

$$\begin{aligned} x_1 &= L_1(Y)=\alpha_{11}y_1+\alpha_{12}y_2+\ldots+\alpha_{1q_1}y_{q_1} \\ x_2 &= L_2(Y)=\alpha_{21}y_1+\alpha_{22}y_2+\ldots+\alpha_{2q_2}x_{q_2} \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}$$

where the linear forms $L_m(Y)$, each containing only a finite number of the coordinates of Y, fulfill the following two conditions. 1°. The $L_m(Y)$ are linearly independent, i. e. there exists no linear relation with constant, not all vanishing coefficients among any finite number of them. 2^0. Each variable y_q ($q = 1, 2, \ldots$) may be "isolated", i. e. expressed as a linear combination with constant coefficients of a finite number of the $L_m(Y)$. For a later application we remark that if a (finite or enumerable) set of linear forms satisfies only the condition 1^0 we may always (III, p. 14) add to the set new linear forms — and even such consisting each of only one coordinate — so that also condition 2^0 be fulfilled.

5. The notion of almost periodicity of a function $F(X)$ on our infinite-dimensional space is invariant under any substitution T performed on X, i. e. $F(TX)$ is again an a. p. function of X. In fact the set of the a. p functions $F(X)$ may be characterized as the closure of the set of the trigonometric polynomials $S(X)$, and a trigonometric polynomial $S(X)$ is evidently transformed into a trigonometric polynomial $S(TX)$. Now, applying a substitution T to a function $P(X)$, p. p. with respect to the axis, and denoting again the new variable point by X, instead of Y, we obtain a (continuous) function which we denote by $P_T(X)$ and abbreviatively call a p. p. function with respect to the substitution T (or more correctly with respect to the straight lines into which the coordinate axis are transformed, and which again span the whole space). For a fixed substitution T and all the $P(X)$ we form the class $\{P_T(X)\}$ and its closure $C_T = \mathrm{Cl}\,\{P_T(X)\}$, the functions of which we call l. p. functions with respect to T. Finally, we form the union Γ of all these classes C_T, $\Gamma = \bigcup_T C_T$, the functions of which we simply denote as l. p. functions.

6. Before proceeding, it may be illustrating to consider the notion of periodicity in our infinite-dimensional space from a more general point of view. A vector $V = (v_1, v_2, \ldots)$ in our space is called a period of the (continuous) function $F(X)$ if $F(X+V) = F(X)$ for all X. Each function has the trivial period $(0, 0, \ldots)$. Obviously, on account of the continuity of $F(X)$ the set of all periods of $F(X)$ is a closed module. Now, according to a simple, but not trivial theorem of E. Følner and myself (IV, p. 30 or V, p. 46) every closed module in our space may be transformed by a substitution T into a module of the simple type $\{(v_1, v_2, \ldots, v_n, \ldots)\}$ where the indices $1, 2, \ldots, n, \ldots$ fall into three classes $\{n_r\}$, $\{n_s\}$, $\{n_t\}$ such that the coordinates v_{n_r} independently run through all real numbers, the coordinates v_{n_s} independently run through all integers while the remaining coordinates v_{n_t} are all equal to 0. Here we are only interested in the case in which the last class $\{n_t\}$ is empty, as otherwise the module does not span the whole space (i. e. is lying in a proper subspace). Thus we see that there exists no other function $F(X)$ with a period module which span the whole space than those introduced above, i. e. functions belonging to one of the classes $\{P_T(X)\}$.

7. Returning to the l. p. functions, there exists in case of functions of infinitely many variables a problem which has no analogue for l. p. functions of one variable[1]) and to which B. JESSEN has called my attention, namely *whether the union Γ of all the closed sets $C_T = \mathrm{Cl}\{P_T(X)\}$ is identical with (or only forms a part of) the closure Γ^* of the whole set of all the p. p. functions.* Since Γ^* is closed and Γ obviously contains all the p. p. functions, the problem is, in other words, whether the set Γ is closed. The purpose of this paper is to give the solution of this problem by proving

Theorem. *The set $\Gamma = \bigcup\limits_T C_T$ consisting of all the l. p. functions of infinitely many variables is a closed one.*

8. The proof to be given in the next section depends on the consideration of the Fourier series of the l. p. functions $G(X) = G(x_1, x_2, \ldots)$. From what have been said before it easily follows that a necessary and sufficient condition for a Fourier series of an a. p. function $F(X)$ to be that of one of our l. p. functions is that the linear forms in the exponents

$$M_n(X) = \Lambda_{n,1}x_1 + \Lambda_{n,2}x_2 + \ldots + \Lambda_{n,\,m_n}x_{m_n}$$

can be obtained from linear forms with mere rational coefficients by subjecting them to some linear substitution. From this characterization of the Fourier series of the l. p. functions we shall deduce the following

Lemma. *A necessary and sufficient condition for a Fourier series $\Sigma A_n e^{iM_n(X)}$ of an a. p. function to belong to an l. p. function is that in any relation which expresses one of the linear forms $M_n(X)$ as a linear combination of a finite number of linearly independent forms of the sequence $M_1(X), M_2(X), \ldots$, the occurring constant coefficients (uniquely determined) shall all be rational.*

That the condition is necessary can immediately be seen. In fact, as a linear substitution does not change linear relations, or linear independance, among the linear forms in the exponents, it suffices to prove that the condition is fulfilled for a function l. p. with respect to the axis, for instance of the special class $C_{2\pi, 2\pi, \ldots}$. But in this case all the occurring coefficients $\Lambda_{n,m}$ are rational numbers and hence the condition is evidently fulfilled since a finite number of ordinary linear equations with rational coefficients and only one solution can only have a solution in rational numbers.

In order to see that the condition is sufficient we proceed in the following manner. From the sequence $M_1(X), M_2(X), \ldots$ we first select (successively) a subsequence of which any finite number of its terms is linearly independent and such that any $M_n(X)$ may be expressed as a linear

[1]) The problem exists also for l. p. functions of a finite number of variables $x_1, x_2, \ldots, x_n$ $(n > 1)$ and the solution given below is also valid in this case. However, in the finite-dimensional case the problem may easily be solved without applying the theory of Fourier series.

combination of a finite number of forms of this subsequence. To the linear forms of this subsequence we may, as remarked before, add new linear forms such that the enlarged (enumerable) set of linear forms may be used as right-hand sides of a linear substitution. By performing this substitution — or rather the inverse one — on our Fourier series $\Sigma A_n e^{iM_n(X)}$ we evidently obtain, from the assumption that the condition be fulfilled, a new Fourier series with mere rational coefficients in the exponents. Thus the corresponding function and hence also the function $F(X)$ before the transformation is a l.p. function.

9. We can now easily prove our theorem, viz. that the set Γ of all l. p. functions $G(X)$ is closed. We have to prove that if $F(X) \sim \Sigma A_n e^{iM_n(X)}$ is an a. p., but not a l.p. function, then $F(X)$ cannot be approximated uniformly by l. p functions. According to our lemma there exists among the linear forms $M_n(X)$ in the exponent of the Fourier series of $F(X)$ a linear relation

$$M_N(X) = b_1 M_{n_1}(X) + b_2 M_{n_2}(X) + \ldots + b_s M_{n_s}(X)$$

with linearly independent $M_{n_1}(X), M_{n_2}(X), \ldots, M_{n_s}(X)$ and not all b's rational. Now, as well-known, uniform convergence of a sequence of a. p. functions towards an (a. p.) function $F(X)$ implies formal convergence of the Fourier series of the functions of the sequence towards the Fourier series $\Sigma A_n e^{iM_n(X)}$ of $F(X)$. Hence in the Fourier series of any a. p. function $H(X)$ which approximates $F(X)$ sufficiently close each of the finite number of linear forms $M_N(X), M_{n_1}(X), \ldots, M_{n_s}(X)$ must necessarily occur as exponents, simply because they occur in the Fourier series of $F(X)$. Consequently, using the lemma once more, we see that the a. p. function $H(X)$ cannot be l. p. This proves the theorem.

References.

I. S. Bochner, Beiträge zur Theorie der fastperiodischen Funktionen. II. Teil, *Math. Annalen*, **96** (1926), pp. 383–409.

II. H. Bohr, Zur Theorie der fastperiodischen Funktionen. II. Teil, *Acta Math.*, **46** (1925), pp. 101—214.

III. H. Bohr, Unendlich viele lineare Kongruenzen mit unendlich vielen Unbekannten, *Kongelige Danske Videnskabernes Selskab, Math.-fys. Meddelelser*, **7** (1926), No. 1.

IV. H. Bohr and E. Følner, Infinite systems of linear congruences with infinitely many variables, *Ibidem*, **24** (1948), No. 12.

V. H. Bohr and E. Følner, On a structure theorem for closed modules in an infinite-dimensional space, *Studies and Essays presented to R. Courant* (New York, 1948), pp. 45–62.

(Received September 21, 1949.)

Remarque sur le prolongement fonctionnel linéaire et le problème de Dirichlet.

Par M. BRELOT à Grenoble (France).

1. Considérons dans l'espace euclidien R_τ de dimension $\tau \geqq 2$ un domaine Ω, par exemple borné, de frontière Ω^*. A toute donnée f finie continue réelle sur Ω^* correspond la solution du problème de Dirichlet généralisé ou fonction de Wiener H_f; et KELDYCH[1]) a montré que c'est la seule fonction harmonique dans Ω qui soit, en chaque point, fonction linéaire homogène croissante de f et qui coïncide avec la solution du problème de Dirichlet classique (tendant vers $f(P)$ en tout point P de Ω^*) lorsque celle ci existe. C'est-à-dire que, si on part de cette solution classique qui définit (avec la notion d'ordre de fonctions correspondant à une même inégalité partout) une application linéaire homogène croissante de l'ensemble $\mathcal{T}$ de certains f dans l'ensemble ordonné $\mathcal{E}$ des fonctions harmoniques réelles dans Ω, il y a une seule manière de faire le prolongement à tous les f en conservant linéarité, homogénéité, et croissance.

Or, si l'on examine le *théorème de Hahn—Banach*[2]) *sur le prolongement fonctionnel* lorsque, l'espace de la variable étant ordonné, la fonction linéaire et la fonction convexe majorante sont croissantes d'où la croissance de la fonction prolongée, on va voir qu'il est aisé de l'adapter à notre application linéaire et que l'unicité, connue à priori, entraine l'identité avec H_f de deux fonctions harmoniques $\underline{D}_f, \overline{D}_f$ que précise l'énoncé suivant:

Théorème 1. *Reprenons f finie continue sur la frontière Ω^* du domaine borné Ω. Les solutions du problème de Dirichlet classique prenant des valeurs-frontière $\leqq f$ ont une enveloppe supérieure finie continue sousharmonique $\underline{\mathfrak{D}}_f$*

---

[1]) M. V. KELDYCH, Sur le problème de Dirichlet, *Comptes Rendus Acad. Sci. URSS*, **32** (1941), p. 308–309. Au lieu de la croissance en f, KELDYCH suppose d'ailleurs, ce qui est équivalent, que la fonction soit comprise (au sens large) entre les bornes de f. Il ne précise pas non plus la nature du domaine et le nombre des dimensions de l'espace, mais nos hypothèses suffisent et il sera discuté de ce point plus loin.

[2]) Voir S. BANACH, *Théorie des opérations linéaires* (Varsovie, 1932), p. 27–29.

dont la plus petite majorante harmonique $\underline{D}_f$ est la borne supérieure des solutions considérées dans l'ensemble ordonné $\mathcal{E}$. De même les solutions $\geqq f$ à la frontière ont une enveloppe inférieure $\overline{\mathfrak{D}}_f$ finie, continue surharmonique dont la plus grande minorante harmonique $\overline{D}_f$ est la borne inferieure de ces solutions. De plus $\underline{D}_f = \overline{D}_f = H_f$.

D'abord toute famille de fonctions harmoniques u dans un domaine, bornées supérieurement dans leur ensemble, a une enveloppe supérieure continue sousharmonique; car ces fonctions sont également continues en chaque point quand on prend sur la droite de leurs valeurs la structure uniforme de la droite numérique achevée, c'est-à-dire que les e^u sont également continues au sens ordinaire[3]). L'enveloppe est donc continue et la sousharmonicité est immédiate. D'où l'existence d'une borne supérieure des u dans l'ensemble ordonné des fonctions harmoniques sur le domaine. Il s'ensuit les résultats concernant séparément $\underline{\mathfrak{D}}_f$, $\underline{D}_f$ et de même $\overline{\mathfrak{D}}_f$, $\overline{D}_f$.

Cherchons maintenant à adapter la démonstration de Banach en remplaçant la fonction convexe homogène par $\overline{D}_f$ qui satisfait visiblement à

$$\overline{D}_{f+g} \leqq \overline{D}_f + \overline{D}_g, \quad \overline{D}_{tf} = t\overline{D}_f \quad (t \geqq 0),$$

$\overline{D}_f$ est croissante de f et majore la fonction linéaire homogène H_f prise dans l'espace vectoriel $\mathcal{T}$.

Quels que soient $f' \in \mathcal{T}$, $f'' \in \mathcal{T}$, $f_0 \overline{\in} \mathcal{T}$ on verra que

$$-\overline{D}_{-f'-f_0} - H_{f'} \leqq \overline{D}_{f''+f_0} - H_{f''}$$

La borne inférieure dans $\mathcal{E}$ du premier membre pour f' variable dans $\mathcal{T}$ et la borne supérieure dans $\mathcal{E}$ du second membre pour f'' variable dans $\mathcal{T}$ sont justement $\underline{D}_{f_0}$ et $\overline{D}_{f_0}$.

Si elles étaient distinctes on pourrait choisir avec arbitraire une fonction harmonique intermédiaire h_0 puis faire un premier prolongement P_φ défini pour les $\varphi = f + tf_0$ ($f \in \mathcal{T}$, t réel) selon $P_\varphi = H_f + th_0$ et dont on voit qu'il est additif, homogène majoré par $\overline{D}_\varphi$.

Le raisonnement final de BANACH s'applique encore et nous fournirait un prolongement pour tous les f, majoré par $\overline{D}_f$ croissant, donc croissant; et l'arbitraire de h_0 correspondant à f_0 serait contraire à l'unicité. On a donc bien $\underline{D}_f = \overline{D}_f$ et comme a priori elles encadrent H_f, le théorème s'ensuit.

2. Ces développements appellent divers commentaires et recherches.

D'abord on peut songer à étendre notre démonstration de prolongement à des espaces vectoriels généraux convenablement ordonnés. C'est ce qu'a fait M. CHOQUET dans des recherches et ce n'est peut-être pas sans rapport avec des résultats récents (à l'impression) de NACHBIN sur le prolongement

---

[3]) Voir M. BRELOT, Sur le rôle du point à l'infini dans la théorie des fonctions harmoniques, *Annales de l'École Normale Supérieure,* **61** (1945), p. 301–332 et particulièrement p. 315. Cela peut se déduire des inégalités de HARNACK.

d'une application linéaire dans un espace normé, d'un sous-espace d'un autre espace normé.

D'autre part on songera à étendre le théorème 1 à un domaine plus général (par exemple dans l'espace compact $\overline{R}_r$, Ω étant seulement de complémentaire non polaire[4]) ou à un problème de Dirichlet plus général du type "ramifié" ou géodésique[5]). Tout dépend des conditions de validité du résultat d'unicité de Keldych, résultat qui est conséquence presque immédiate de la propriété suivante[6]) des points réguliers donnée aussi par KELDYCH[7]):

(K): Si 0 est point-frontière régulier de Ω borné, il existe f finie continue sur Ω atteignant son minimum au seul point 0 et pour laquelle le problème de Dirichlet classique admet une solution.

C'est cette propriété qu'il faudrait étendre. Notons seulement qu'elle entraine, soit directement, soit par l'intermédiaire du théorème 1, et cela ne parait pas avoir été encore explicité, que la solution généralisée H_f (f finie continue) est l'enveloppe supérieure des fonctions finies continues dans $\Omega \cup \Omega^*$ sousharmoniques dans Ω majorées par f sur Ω^*[8])

Enfin on peut se demander avec M. CHOQUET, si plutôt et mieux que ce dernier énoncé, on n'aurait pas $H_f = \mathfrak{D}_f$. Il est aisé de voir que c'est inexact, avec un seul point irrégulier qui serait point-frontière isolé (Ω cercle pointé), mais le résultat parait probable lorsque la frontière est *débarrassée de sa partie impropre*, (c'est-à-dire est de capacité >0 au voisinage de chacun de ses points); on peut déjà s'en assurer lorsqu'il n'y a qu'un nombre fini de points irréguliers restants.

(Reçu le 19 octobre 1949)

---

4) Voir le mémoire cité note 3).

5) Voir M. BRELOT, Le problème de Dirichlet géodésique, *Comptes Rendus Acad. Sci. Paris*, **228** (1949), p. 1790–1792.

6) M. V. KELDYCH en déduit en effet que son opérateur est dans Ω une fonction harmonique bornée qui tend vers $f(P)$ en tout point-frontière régulier P, ce qui l'identifie à H_f.

7) M. V. KELDYCH, On the solubility and the stability of Dirichlet problem, *Uspechi Mat. Nauk*, **8** (1941), p. 171–231 et spécialement p. 226. (Russian). L'auteur se place dans l'espace ordinaire mais le type de démonstration qui s'applique aux espaces supérieurs s'adapte au cas du plan.

8) Cette famille de fonctions sousharmoniques est celle considérée à l'origine par O. PERRON, Eine neue Behandlung der ersten Randwertaufgabe für $\Delta u = 0$, *Math. Zeitschrift*, **18** (1923), p. 42–54, avant qu'on ne considère la famille plus étendue des fonctions sousharmoniques sur Ω et qui, à la fontière ont seulement une lim sup majorée par f. Dans ce même mémoire, PERRON, cherchant des critères de résolubilité du problème classique, introduit sur les points-frontière une condition restrictive dont on peut voir seulement maintenant grâce à (K) qu'il s'agit exactement de la régularité.

Commutativity and spectral properties of normal operators.

By PAUL R. HALMOS in Chicago.

1. The results of this note grew out of a current investigation of spectral properties of operators on Hilbert space. While the characterization of the spectral manifolds of a normal operator (Theorem 1) appears to be new and may be considered to be of independent interest, I present it here mainly because it supplies an extremely easy proof of a theorem (Theorem 2) which was unknown until a few weeks ago. J. VON NEUMANN has asked whether or not it is true that if an operator B commutes with a normal operator A, then B commutes with A^* also. Well known and quite elementary considerations show that in order to answer the question affirmatively it is sufficient to prove that, under the stated hypotheses, B is reduced by all the spectral manifolds of A. This has recently been proved by B. FUGLEDE — he communicated his proof to me at the Boulder meeting of the American Mathematical Society at the end of August, 1949. The proof I present below is somewhat different from his in spirit and in method. I should say also that FUGLEDE's proof is valid for not necessarily bounded transformations A and that, similarly, only minor modifications are needed to adapt my proof to this more general case.

For the orientation of the reader I present here the trivial proof of the theorem under discussion for the case in which A has pure point spectrum; the proof of the general theorem below uses essentially the same idea and method. If λ is a proper value of A and if $\mathfrak{F}$ is the subspace of all corresponding proper vectors, then the relations $A(Bx)=B(Ax)=B(\lambda x)=\lambda(Bx)$ show that $\mathfrak{F}$ is invariant under B. Since to say that A has pure point spectrum means that the entire Hilbert space is spanned by orthogonal subspaces such as $\mathfrak{F}$, it follows that the orthogonal complement of $\mathfrak{F}$ is also invariant under B, and this is exactly what was to be proved.

2. Throughout this note I shall deal with a fixed complex Hilbert space $\mathfrak{H}$. An *operator* is a bounded linear transformation of $\mathfrak{H}$ into itself; an operator A is *normal* if it commutes with its adjoint A^*. If A is normal, then

$\|Ax\|^2 = (Ax, Ax) = (A^*Ax, x) = (AA^*x, x) = (A^*x, A^*x) = \|A^*x\|^2$ for every vector x; it is easy to see that the identity $\|Ax\| = \|A^*x\|$ is not only necessary but also sufficient for the normality of A. A *subspace* is a closed linear manifold in $\mathfrak{H}$; a subspace $\mathfrak{M}$ *reduces* an operator A if both $\mathfrak{M}$ and $\mathfrak{M}^\perp$ (= the orthogonal complement of $\mathfrak{M}$) are invariant under A, i. e. if $A\mathfrak{M} \subset \mathfrak{M}$ and $A\mathfrak{M}^\perp \subset \mathfrak{M}^\perp$. There are two useful and elementary necessary and sufficient conditions that a subspace $\mathfrak{M}$ reduce an operator A; the first is that $\mathfrak{M}$ be invariant under both A and A^*, and the second is that A commute with the projection on $\mathfrak{M}$.

Lemma 1.[1]) *If A is a normal operator and if $\mathfrak{F}(A) = \{x: \|A^n x\| \leqq \|x\|, n = 1, 2, \ldots\}$, then $\mathfrak{F}(A)$ is a subspace and $\mathfrak{F}(A)$ is invariant under every operator B which commutes with A.*

Proof. Write $\mathfrak{S}$ for the set of all those vectors x for which the sequence $\{\|A^n x\| : n = 1, 2, \ldots\}$ is bounded. Since $\|A^n(\alpha x)\| = |\alpha| \cdot \|A^n x\|$ and $\|A^n(x+y)\| \leqq \|A^n x\| + \|A^n y\|$, it follows that $\mathfrak{S}$ is a linear manifold; if an operator B commutes with A, then the relation $\|A^n(Bx)\| = \|B(A^n x)\| \leqq \leqq \|B\| \cdot \|A^n x\|$ implies that $\mathfrak{S}$ is invariant under B. Clearly $\mathfrak{F}(A)$ is a closed set and $\mathfrak{F}(A) \subset \mathfrak{S}$; the proof of the lemma will be completed by showing that $\mathfrak{F}(A) = \mathfrak{S}$. For this purpose it is sufficient to show that if x is a vector such that, for some positive integer p, $\|A^p x\| > \alpha \|x\|$, $\alpha > 1$, then the sequence $\{\|A^n x\|\}$ cannot be bounded. Since $\alpha^2 \|x\|^2 < \|A^p x\|^2 = (A^p x, A^p x) = (A^{*p} A^p x, x) \leqq \leqq \|A^{*p} A^p x\| \cdot \|x\| = \|A^{2p} x\| \cdot \|x\|$, it follows that $\|A^{2p} x\| > \alpha^2 \|x\|$. Since an inductive repetition of this argument shows that $\|A^{2^k} x\| > \alpha^{2^k} \|x\|$ for every positive integer k, the proof is complete.

3. A *spectral measure* is a function E from the class of all Borel subsets of the set Λ of all complex numbers to projections on $\mathfrak{H}$, such that (i) $E(\Lambda) = 1$, (ii) $E(M \cap N) = E(M) E(N)$ whenever M and N are Borel sets, and (iii) $E(M) = \sum_{j=1}^{\infty} E(M_j)$ whenever $\{M_j\}$ is a disjoint sequence of Borel sets whose union is M (the series being understood to converge in the strong topology of operators).

Lemma 2. *If E is a spectral measure and if $\mathfrak{E}(M) = \{x: E(M)x = x\}$ for every Borel set M, then $\mathfrak{E}(M)$ is the subspace spanned by the class of all subspaces of the form $\mathfrak{E}(N)$, where N is an arbitrary compact subset of M.*

Proof. The assertion of the theorem is that, in a sense well known in the theory of numerical measures, every spectral measure is regular. The proof may be given along lines entirely similar to the numerical case, or it

[1]) This lemma is proved for Hermitian operators by B. A. Lengyel and M. H. Stone, Elementary proof of the spectral theorem, *Annals of Math.*, **37** (1936), pp. 853—864; cf. in particular p. 858. The following proof is a slight simplification of their proof.

may be reduced to that case as follows. All that it is necessary to prove is that if x is a vector in $\mathfrak{E}(M)$ such that x is orthogonal to $\mathfrak{E}(N)$ for every compact subset N of M, then $x=0$. Since, however, by the regularity of numerical measures, $\|x\|^2=\|E(M)x\|^2=\sup_N \|E(N)x\|^2$, it follows that there exists a countable class $\{N_j\}$ of compact subsets of M such that $\|x\|^2=\sup_j \|E(N_j)x\|^2$, and hence that indeed $x=0$.

I shall make use below of the spectral theorem for normal operators in the following form. If A is a normal operator, then there exists a unique spectral measure E, called the spectral measure of A, such that $(Ax, y)=$ $=\int \lambda\, d(E(\lambda)x, y)$ for every pair of vectors x and y.

4. In this final section I shall assume that A is a fixed normal operator with spectral measure E. For every complex number λ and every positive real number ε, I shall write $\mathfrak{F}(\lambda, \varepsilon)$ for $\mathfrak{F}\left(\frac{A-\lambda}{\varepsilon}\right)$; for every set M of complex numbers and every positive real number ε, I shall write $\mathfrak{F}(M, \varepsilon)$ for the subspace spanned by all those $\mathfrak{F}(\lambda, \varepsilon)$ for which $\lambda \in M$; and, for every set M of complex numbers, I shall write $\mathfrak{F}(M)=\bigcap_{\varepsilon>0} \mathfrak{F}(M, \varepsilon)$. Let $F(\lambda, \varepsilon)$, $F(M, \varepsilon)$, and $F(M)$ be the projections on the subspace $\mathfrak{F}(\lambda, \varepsilon)$, $\mathfrak{F}(M, \varepsilon)$, and $\mathfrak{F}(M)$, respectively.

Theorem 1. *For every compact set M, $\mathfrak{F}(M)=\mathfrak{E}(M)$.*

Proof. For any positive number ε, let $\{M_j\}$ be a disjoint sequence of non empty Borel sets of diameter not greater than ε and such that $\bigcup_j M_j=M$. If $x \in \mathfrak{E}(M)$, $x_j=E(M_j)x$, and $\lambda_j \in M_j$, then $\|(A-\lambda_j)^n x_j\|^2=$ $=\int_{M_j} |(\lambda-\lambda_j)^n|^2\, d(E(\lambda)x_j, x_j) \leqq \varepsilon^{2n}\|x_j\|^2$, so that, for each j, $x_j \in \mathfrak{F}(\lambda_j, \varepsilon) \subset \mathfrak{F}(M, \varepsilon)$. Since $x=E(M)x=\Sigma_j E(M_j)x=\Sigma_j x_j$, it follows that $x \in \mathfrak{F}(M, \varepsilon)$. The arbitrariness of ε implies that $x \in \mathfrak{F}(M)$, and the arbitrariness of x implies, consequently, that $\mathfrak{E}(M) \subset \mathfrak{F}(M)$. Note that this argument did not make use of compactness of M.

Suppose now that N is a compact subset of $\Lambda - M$, and let δ be the distance between M and N. If $\lambda_0 \in M$, if $0<\varepsilon<\delta$, and if $x \in \mathfrak{F}(\lambda_0, \varepsilon)$, then $\|(A-\lambda_0)^n x\| \leqq \varepsilon^n \|x\|$; if, on the other hand, $x \in \mathfrak{E}(N)$, then $\|(A-\lambda_0)^n x\|^2=$ $=\int_N |(\lambda-\lambda_0)^n|^2\, d(E(\lambda)x, x) \geqq \delta^{2n}\|x\|^2$. It follows that $\mathfrak{F}(\lambda_0, \varepsilon) \cap \mathfrak{E}(N)=\{0\}$. Since $E(N)$ commutes with A, it follows from Lemma 1 that $\mathfrak{F}(\lambda_0, \varepsilon)$ is invariant under $E(N)$ and hence, since $E(N)$ is Hermitian, that $E(N)$ commutes with $F(\lambda_0, \varepsilon)$. This in turn implies that $F(\lambda_0, \varepsilon)E(N)$ is the projection on $\mathfrak{F}(\lambda_0, \varepsilon) \cap \mathfrak{E}(N)$, i. e. that $F(\lambda_0, \varepsilon)E(N)=0$, and it follows that $\mathfrak{F}(\lambda_0, \varepsilon)$ is orthogonal to $\mathfrak{E}(N)$. The validity of this assertion for every λ_0 in M shows that $\mathfrak{F}(M, \varepsilon)$ is orthogonal to $\mathfrak{E}(N)$ and therefore, *a fortiori*, that $\mathfrak{F}(M)$ is orthogonal to $\mathfrak{E}(N)$.

The result of the preceding paragraph implies, in view of Lemma 2, that $\mathfrak{F}(M)$ is orthogonal to $\mathfrak{E}(\Lambda - M)$. This means that $\mathfrak{F}(M) \subset (\mathfrak{E}(\Lambda - M))^{\perp} = \mathfrak{E}(M)$, and the proof of the theorem is complete. I remark that it is easy to construct examples to show that if M is not compact, then $\mathfrak{E}(M)$ may be a proper subset of $\mathfrak{F}(M)$.

Theorem 2. *If an operator B commutes with A, then $\mathfrak{E}(M)$ reduces B for every Borel set M.*

Proof. It follows from Lemma 1 that, for every complex number λ and every positive number ε, $\mathfrak{F}(\lambda, \varepsilon)$ is invariant under B, and hence that $\mathfrak{F}(M, \varepsilon)$ and $\mathfrak{F}(M)$ are invariant under B for every set M. Theorem 1 implies that $\mathfrak{E}(M)$ is invariant under B whenever M is compact and hence, by Lemma 2, that $\mathfrak{E}(M)$ is invariant under B for every Borel set M. Since $(\mathfrak{E}(M))^{\perp} = \mathfrak{E}(\Lambda - M)$, it follows automatically that $(\mathfrak{E}(M))^{\perp}$ is also invariant under B and hence that $\mathfrak{E}(M)$ reduces B.

University of Chicago.

(Received October 13, 1949.)

The class of functions which are absolutely convergent Fourier transforms.

By I. E. SEGAL in Chicago.

We show that if a locally compact abelian group G has the property that every continuous function vanishing at infinity on it is an absolutely convergent Fourier transform, then G is finite. More precisely, the set of functions on G which are absolutely convergent Fourier transforms is dense (in the uniform topology) in the space of continuous functions on G which vanish at infinity, and is either all of the space, or of first category in it, according as G is finite or not. Thus the difficulties in Fourier analysis which appear to arise from the circumstance that not every continuous function vanishing at infinity on a locally compact abelian group is an absolutely convergent Fourier transform are inherent in any generalization from finite groups. A point of interest from a more classical viewpoint is that our results establish the existence of continuous functions on the reals which tend to zero at infinity, but which are not absolutely convergent Fourier transforms, and also the existence of similar continuous periodic functions on the reals. These results are well known, by virtue of explicit constructions of such functions, but the proof which we give is purely existential, and thereby avoids the computations required to verify that given functions are not representable as absolutely convergent Fourier transforms. (BANACH has shown in a related fashion the existence of nondifferentiable functions.)

We are indebted to W. AMBROSE, I. KAPLANSKY and D. MONTGOMERY for very helpful conversations.

Theorem. *Let G be a locally compact abelian (abbreviated to: "LCA") group. Then the class of all functions on G which are absolutely convergent Fourier transforms is either a dense set of first category in the space of complex-valued continuous functions vanishing at infinity on G, or all of that space, according as G is infinite or finite.*

We recall that a function f on G is said to be an absolutely convergent Fourier transform if there is a function F, integrable on the character group G^* of G, such that $f(x)=\int_{G^*} x^*(x)\,F(x^*)\,dx^*$, for $x\in G$ (here and elsewhere dx^* represents the element of Haar measure on G^*, dx that on G, and "integrable" means "integrable relative to Haar measure"). A function φ vanishes at infinity on a topological space Γ if for every positive number ε, the set where φ exceeds ε in absolute value has compact closure. By $C(\Gamma)$ we denote the Banach algebra off all complexvalued continuous functions vanishing at infinity on the space Γ, the norm of a function being defined as its least upper bound. A statement will be said to be true "nearly everywhere" (abbreviated to "n. e.") on a LCA group G if it is a function on the group which is true except on a set whose intersection with any compact set has Haar measure zero, and a numerical function on G will be called "measurable" if the inverse image of any open set under the function is a set in G which meets each compact set in a set which is measurable in the usual sense. For any LCA group G, G^* will denote the character group of G, and $L_p(G)$ will denote the Banach space of pth-power integrable complexvalued functions on G, the norm of a function being defined as the pth root of the integral of the pth power of the absolute value of the function.

Lemma 1. *If K is a regular finite countably-additive set function on a LCA group G, then its Fourier-Stieltjes transform $\int_G x^*(x)\,dK(x)$ vanishes identically (on G^*) only if $K=0$.*

If h is an arbitrary continuous function on G which vanishes outside of a compact set, then $\int h(xy^{-1})\,dK(y)$ defines a function on G which is easily shown, by means of the Fubini theorem, to be integrable, and whose Fourier transform is the product of the Fourier-Stieltjes transform of K with the Fourier transform of h, and so vanishes identically. By the uniqueness theorem for Fourier transforms (see e. g. [3]), $\int_G h(xy^{-1})\,dK(y)$ vanishes n. e., so that being a continuous function, it vanishes everywhere on G. By virtue of the arbitrary character of h and the regularity of K, it results that $K=0$.

The following lemma is known, but the simple proof which we give is new.

Lemma 2. *For any LCA group G, the Fourier transforms of the elements of $L_1(G)$ are dense in $C(G^*)$.*

It is known (the generalized Riemann-Lebesgue lemma) that all L_1 transforms are in $C(G^*)$ (see e. g. [4]), and we prove their density by an indirect argument. Assuming that they are not dense, then by the Hahn-Banach theorem there exists a non-vanishing continuous linear functional χ on $C(G^*)$ which vanishes on all transforms. By the known form of continuous linear functionals on $C(\Gamma)$, where Γ is a locally compact Hausdorff space, there exists a finite regular countably-additive set function K on G^* such that for $F\in C(G^*)$, $\chi(F)=\int_{G^*} F(x^*)\,dK(x^*)$. Now if $F(x^*)=\int_G x^*(x)f(x)\,dx$, with

$f \in L_1(G)$, then $\chi(F) = 0 = \int\limits_{G^*} \left[\int\limits_{G} x^*(x) f(x)\, dx \right] dK(x^*)$, which by the Fubini theorem is the same as $\int\limits_{G} \left[\int\limits_{G^*} x^*(x)\, dK(x^*) \right] f(x)\, dx$. Now f is arbitrary in $L_1(G)$, so that it follows that $\int\limits_{G^*} x^*(x)\, dK(x^*) = 0$ n. e. on G, and since the integral represents a continuous function, vanishes identically on G. Lemma 1 now implies $K = 0$, which yields the contradiction $\chi = 0$.

Lemma 3. *If T is an open subgroup of a LCA group G, then an element of $L_1(G)$ vanishes n. e. outside of T if and only if its Fourier transform is constant on cosets modulo the annihilator $\widetilde{T}$ of T in G^*.*

It is trivial to verify the "only if" part. Now if F is the Fourier transform of an element f of $L_1(G)$, and is constant on cosets modulo $\widetilde{T}$, then $F(x^*y^*) = F(x^*)$ for all $y^* \in \widetilde{T}$ and $x^* \in G^*$. It results that $\int\limits_{G} x^*(x)(y^*(x) - 1) f(x)\, dx = 0$ for $x^* \in G^*$, which by the uniqueness of the Fourier transform implies that $[y^*(x) - 1] f(x)$ vanishes n. e. If f is continuous, this implies in turn that f vanishes except on the set of x for which $y^*(x) = 1$, and since y^* is arbitrary on $\widetilde{T}$, it follows that f then vanishes except on the intersection over $y^* \in \widetilde{T}$ of such sets, i. e. f vanishes outside of T. Now in general, there exists a sequence $\{g_n\}$ of bounded functions in $L_1(G)$ which vanish outside of T and are such that $\{f * g_n\}$ converges in $L_1(G)$ to f. The continuous function $f * g_n$ has as its transform the product of the Fourier transforms of f and g_n and so is constant on cosets modulo $\widetilde{T}$. Hence $f * g_n$ vanishes outside of T, and since the set of elements of $L_1(G)$ which vanish n. e. outside T is closed, f itself is in this set.

It is convenient at this point to introduce the following notation: a LCA group G has the property Φ, or alternatively G^* has property Φ^*, if every element of $C(G)$ is an L_1 Fourier transform. We note that when G has the property Φ, then the Fourier transform is a homeomorphism of $L_1(G^*)$ onto $C(G)$, by a well-known theorem of BANACH [1].

Lemma 4. *If T^* is an open subgroup of the LCA group G^*, and if G^* has the property Φ^*, then so also does T^*.*

If $f \in C(T^{**})$, then there is a unique function f' in $C(G/\widetilde{T}^*)$ which corresponds to f via the natural isomorphism of T^{**} with $G/\widetilde{T}^*$, and this function defines via the inverse of the natural mapping on G to $G/\widetilde{T}^*$, an element f'' of $C(G)$ which is constant on cosets modulo $\widetilde{T}^*$. By the preceding lemma there is an element F'' of $L_1(G^*)$ which vanishes outside T^* and is such that

$f''(x)=\int_{T^*}x^*(x)F''(x^*)\,dx^*$. It follows that $f'(x\widetilde{T}^*)=\int_{T^*}x^*(x)F''(x)\,dx^*$ and hence that $f(u)=\int_{T^*}x^*(u)F''(x^*)\,dx^*$ for $u\in T^{**}$.

Lemma 5. *If a compact metric abelian group G has the property Φ, then G is finite.*

Clearly G^* is either a finite or a countable discrete group, and so weak sequential convergence in $L_1(G^*)$ coincides with strong convergence (see [1]. Hence the same is true in $C(G)$. Since weak sequential convergence in a $C(\Gamma)$, where Γ is compact, is identical with convergence at every point, together with uniform boundedness (this fact follows easily from the known form of continuous linear functionals on a $C(\Gamma)$), it follows that any bounded sequence of elements of $C(G)$ which converges at every point converges uniformly. Since the characteristic function of a closed set in G is the bounded pointwise limit of elements of $C(G)$, it results that every such function is continuous, so that every closed set in G is open, i. e. G is discrete. Being also compact, it is finite.

Lemma 6. *If a compact abelian group G has the property Φ, then G is finite.*

For then G^* is discrete, and is either finite, in which case G is also, or contains a countable subgroup H^*. In the latter event, Lemma 4 would imply that H^* has the property Φ^*, and so by the preceding lemma is finite, a contradiction.

Lemma 7. *If a LCA group G which is generated by a compact neighborhood of the identity has property Φ, then G is finite.*

Evidently any continuous linear functional on $L_1(G^*)$ induces via the isomorphism between $L_1(G^*)$ and $C(G)$ a continuous linear functional on $C(G)$. By the known form for such functionals, it results that for any bounded measurable function k on G^*, there exists a finite bounded regular countably-additive set function K on G^* such that $\int k(x^*)f(x^*)\,dx^*=\int F(x)\,dK(x)$ for all $f\in L_1(G^*)$, where F is the Fourier transform of f (we note that G is σ-finite, so that the Riesz representation theorem for continuous linear functionals on L_1 is valid). It is easy to conclude as in the proof of Lemma 2 that $k(x^*)=\int x^*(x)\,dK(x)$ n. e. on G^*, and hence everywhere on G^*, if k is continuous. From this we conclude that every continuous almost periodic (c. a. p.) function on G^* has an absolutely convergent Fourier series. It suffices to show that if $p(x^*)=\int x^*(x)\,dJ(x)$ is a c. a. p. function on G^*, where J has the same properties as K, and also vanishes on points, then $p=0$. Now according to a result of Lyubarskiĭ [2] (stated for connected groups, but whose proof is valid for groups generated by a compact neighborhood of the identity), there exists a sequence $\{C_n^*\}$ of compact subsets of G^* such that the von Neumann mean $M(q)$, for any c. a. p. function q on G^*, is given by the equation $M(q)=\lim_n m_n^{-1}\int_{C_n^*} q(x^*)\,dx^*$, where $m_n=\int_{C_n^*} dx^*$. It results that for any $y\in G$, $M[x^*(y)p(x^*)]=\int_G m_n^{-1}\Big[\int_{C_n^*} x^*(xy)\,dx^*\Big]dJ(x)$. Since $m_n^{-1}\int_{C_n^*} x^*(xy)\,dx^*$ converges boundedly to $M[x^*(xy)]$, which is 0 or 1 according as $x\neq y^{-1}$ or $x=y$, and since J vanishes on points, it follows that $M[x^*(y)\,p(x^*)]=0$. As this is true for all y, $p=0$.

Thus every c. a. p. function on G^* has an absolutely convergent Fourier series. Now it is known that the set of c. a. p. functions on G^* is isomorphic to the set of all continuous functions on a compact group G' which contains a subgroup algebraically isomorphic with G^*, and in such a way as to preserve Fourier series (see e. g. [3]). Then G' is a compact group with property Φ, and hence is finite, from which it follows that G is finite.

Completion of the proof of the theorem. By the theorem of BANACH quoted earlier, either G has property Φ, or the set of Fourier transforms is of first category in $C(G)$; and in any case dense in $C(G)$, by Lemma 2. Now if G has property Φ, let H^* be the open subgroup of G^* generated by a compact neighborhood of the identity in G^*. Then H^* has property Φ^* and so by the preceding lemma in finite. It follows that G^* is discrete, so that G is compact, and hence finite by Lemma 6.

We mention that the foregoing theorem has an analog for non-commutative locally compact groups G: Let $L_1(G)$ be complete relative to the norm $\|f\|'=\|T_f\|$, where $f\in L_1(G)$ and T_f is the operator on $L_2(G)$ given by the equation $T_f g=f*g$, $g\in L_2(G)$; then G is finite. The validity of this analog was recently established by the author in collaboration with I. KAPLANSKY. A different type of analog, whose validity is an open question, is as follows: If G is a LCA group and $1<p<2$, then the Fourier transform maps $L_p(G)$ into a dense subset of $L_q(G^*)$, where $p^{-1}+q^{-1}=1$, which is of second category only if G is finite. The theorem proved here corresponds to the limiting case $p=1$. We note finally that the separable case of our theorem can be established much more briefly than the general case.

References

1. S. BANACH, *Théorie des opérations linéaires* (Warsaw, 1932).
2. G. YA. LYUBARSKIĬ, On the integration in the mean of almost periodic functions on topological groups, *Uspehi Mat. Nauk*, N. S. **3** (1948), pp. 195—201.
3. I. E. SEGAL, The group algebra of a locally compact group, *Transactions American Math. Society*, **61** (1947), pp. 69–105.
4. A. WEIL, *L'intégration dans les groupes topologiques et ses applications* (Paris, 1940).

(Received September 10, 1949.)

A new proof of the general ergodic theorem.

By YAEL NAIM DOWKER in Princeton, N. J.

The purpose of this note is to give a short and elementary proof of HUREWICZ's ergodic theorem [5] (the ergodic theorem without invariant measure). Our proof is a modification of a proof of BIRKHOFF's ergodic theorem [1] given by R. SALEM in his course at M. I. T. SALEM's proof, as well as a proof given by E. HOPF [4], is in turn a modification of one given by H. R. PITT [8]. All the above mentioned proofs are closely related to RIESZ's proof [9] of BIRKHOFF's ergodic theorem. With the aid of the same method we shall also prove a generalization of HUREWICZ's theorem for transformations which are single valued but not necessarily one to one.

Let $(S, \mathfrak{B}, m)$ be a measure space where S is a set of elements called points and denoted by $x, y, \ldots$, $\mathfrak{B}$ a Borel field of subsets of S and m a countably additive non-negative set function defined for the sets belonging to $\mathfrak{B}$. The sets belonging to $\mathfrak{B}$ are called measurable sets and the set function m is called a measure. We assume that $S \in \mathfrak{B}$ and that S is a union of a countable number of measurable sets of finite measure.

Let T be a one-to-one point transformation of S onto itself. We shall say that T is *measurable* if both T and T^{-1} transform measurable sets into measurable sets. Thus if T is measurable so is T^n for $n = 0, \pm, \pm 2, \ldots$. We shall say that T is positively (negatively) *non-singular* if $T(T^{-1})$ transforms sets of measure zero into sets of measure zero. Thus if T is positively non-singular, so is T^n for $n = 2, 3, \ldots$.

Consider now the set functions m_n defined by $m_n(A) = m(T^n A)$, where $A \in \mathfrak{B}$ and $n = 0, 1, 2, \ldots$. If T is measurable and one-to-one, m_n is a countably additive non-negative set function defined for all sets A belonging to $\mathfrak{B}$. If, moreover T^n is positively non-singular, m_n is absolutely continuous with respect to m for $n = 0, 1, 2, \ldots$. It follows then by the Radon—Nikodym theorem that there exists a measurable function $w_n(x)$ such that for every $A \in \mathfrak{B}$ we have $m_n(A) = \int\limits_A w_n(x)\,dm$, $n = 0, 1, 2, \ldots$. By considering approximating sums to the integrals in question one can show that

$$(1) \qquad \int\limits_S f(x)\,dm = \int\limits_S f(T^n x)\,w_n(x)\,dm$$

for $n = 0, 1, 2, \ldots$, for any measurable set A and for any measurable function

$f(x)$ such that either its positive or negative part is integrable. For any $A \in \mathfrak{B}$ consider $m_{i+j}(A) = m(T^{i+j}A) = \int_A w_{i+j}(x)\,dm$. Also

$$m(T^{i+j}A) = m_i(T^jA) = \int_{T^jA} w_i(x)\,dm = \int_A w_i(T^jx)\,w\,(x)\,dm.$$

It follows that

(2) $$w_{i+j}(x) = w_i(T^jx)\,w_j(x) \quad \text{almost everywhere on } S,$$

and it can be assumed with no loss of generality that the equality in (2) holds everywhere on S, for any $i, j = 0, 1, 2, \ldots$.

Consider now any integrable real-valued function $q(x)$ and let

$$q^n(x) = q(x)\,w_0(x) + q(Tx)\,w_1(x) + \ldots + q(T^{n-1}x)\,w_{n-1}(x) = \sum_{i=0}^{n-1} q(T^ix)\,w_i(x).$$

We shall now state and prove HUREWICZ's ergodic theorem in a form given to it by HALMOS [2]. (For the relation between HUREWICZ's theorem and Theorem 1, see [7].)

Theorem I. *If T is a measurable, positively non-singular, one-to-one transformation of S onto itself, if $f(x)$ is integrable and if $h(x)$ is non-negative and such that* $\lim h^n(x) = \infty$ *almost everywhere, then $f^n(x)/h^n(x)$ converges almost everywhere to a finite limit.*

Remark. Our proof of theorem I. like that of HALMOS [2] and the corresponding proofs of HUREWICZ [5], KHINTCHINE [6] and HOPF [3], depends essentially on the following inequality:

Lemma 1. *Let $q(x)$ be any measurable function such that either its positive or negative part is integrable. Let E be the set of points x such that $q^n(x) \geqq 0$ for some n. Then* $\int_E q(x)\,dm \geqq 0$

In fact, the difference between our proof and those mentioned above lies essentially only in the proof of the inequality. We shall therefore restrict ourselves to the proof of Lemma 1.

Lemma 2. *Let $u_0, u_1, \ldots$ be an infinite sequence of real numbers, and N a fixed positive integer. Suppose that*

$$\max_{1 \leqq n \leqq N} \sum_{i=0}^{n-1} u_{i+j} \geqq 0 \qquad \textit{for all } j \geqq 0.$$

Then

$$\sum_{i=0}^{\nu-1} u_i + \sum_{i=\nu}^{\nu+N-1} (u_i)^+ \geqq 0 \qquad \textit{for all } \nu \geqq 1,$$

where $(u_i)^+ = \max(u_i, 0)$.

Proof of the Lemma 2. By assumption, there exists an increasing sequence of integers $n_0 = 0, n_1, n_2, \ldots$ such that $0 < n_k - n_{k-1} \leqq N$, $u_{n_{k-1}} + \ldots + u_{n_k - 1} \geqq 0$, $k = 1, 2, \ldots$. For any ν, let p be such that $n_{p-1} < \nu \leqq n_p$.

Then $n_p < \nu + N$ and consequently

$$\sum_{i=0}^{\nu-1} u_i + \sum_{i=\nu}^{\nu+N-1} (u_i)^+ \geqq \sum_{i=0}^{n_p-1} u_i + \sum_{i=n_p}^{\nu+N-1} (u_i)^+ \geqq \sum_{k=1}^{p} (u_{n_{k-1}} + \dots + u_{n_k-1}) + \sum_{i=n_p}^{\nu+N-1} (u_i)^+ \geqq 0.$$

Proof of Lemma 1. Let $q(x)$ be any measurable function such that either its positive or negative part is integrable. Fix a positive integer N, and let E_N be the set of points x where $q^n(x) \geqq 0$ for some n, $1 \leqq n \leqq N$. Since it is clear that $E_N \subset E_{N+1}$ and that the union of all the sets E_N $(N=1, 2, \dots)$ is equal to E, it suffices to prove that

$$\int_{E_N} q(x)\, dm \geqq 0.$$

Let us put $g(x) = q(x)$ if $x \in E_N$ and $g(x) = 0$ if $x \in S - E_N$. We first notice that $\int_S g(x)\, dm = \int_{E_N} q(x)\, dm$. Thus it suffices to show that $\int_S g(x)\, dm \geqq 0$. Next we notice that $g(x) \geqq q(x)$ for all $x \in S$. This is clear if $x \in E_N$, and $g(x) = 0 > \operatorname*{Max}_{1 \leqq n \leqq N} q^n(x) \geqq q^1(x) = q(x)$ if $x \in S - E_N$. From this follows that

$$\operatorname*{Max}_{1 \leqq n \leqq N} \sum_{i=0}^{n-1} g(T^i x) w_i(x) \equiv \operatorname*{Max}_{1 \leqq n \leqq N} g^n(x) \geqq 0 \text{ for all } x \in S. \tag{3}$$

In fact, $\operatorname*{Max}_{1 \leqq n \leqq N} g^n(x) \geqq \operatorname*{Max}_{1 \leqq n \leqq N} q^n(x) \geqq 0$ if $x \in E_N$ and $\operatorname*{Max}_{1 \leqq n \leqq N} g^n(x) \geqq g^1(x) = g(x) = 0$ if $x \in S - E_N$.

If we replace x by $T^j(x)$ in (3) and multiply both sides by $w_j(x)$, we obtain from (3) and (2) that

$$\operatorname*{Max}_{1 \leqq n \leqq N} \sum_{i=0}^{n-1} g(T^{i+j} x)\, w_{i+j}(x) \geqq 0 \quad \text{for all } x \in S \text{ and all } j \geqq 0.$$

We can thus apply the lemma to the sequence $u_i = g(T^i x)\, w_i(x)$ and obtain

$$G_\nu(x) \equiv \sum_{i=0}^{\nu-1} g(T^i x)\, w_i(x) + \sum_{i=\nu}^{\nu+N-1} (g(T^i x))^+ w_i(x) \geqq 0$$

for all $x \in S$ and all $\nu \geqq 1$. Hence $\int_S G_\nu(x)\, dm \geqq 0$. But from (1) we see that $\int_S G_\nu(x)\, dm = \nu \int_S g(x)\, dm + N \int_S (g(x))^+ dm$. Hence

$$\int_S g(x)\, dm + \frac{N}{\nu} \int_S (g(x))^+ dm \geqq 0.$$

Now if $(g(x))^+$ is integrable, we see by letting ν tend to ∞ that $\int_S g(x)\, dm \geqq 0$. If $(g(x))^+$ is not integrable, then $(g(x))^- = \operatorname{Max}(-g(x), 0)$ is integrable and $\int_S g(x)\, dm = \infty > 0$. (Notice that while in the course of the proof we have used the fact that T is only positively non-singular, it is true that the assumption that T is so together with the assumption that $\lim h^n(x)=\infty$ almost everywhere implies that T is also negatively non-singular.)

We turn our attention now to transformations which are single-valued but not necessarily one-to-one. We will state and prove a generalization of Theorem I. Let $(S, \mathfrak{B}, m)$ be a measure space and let T be a single-valued transformation of S onto itself. We assume that T is measurable, i. e that both TA and $T^{-1}A$ are measurable if A is measurable. We also assume that T is positively non-singular, i. e. that $m(T^{-1}A)=0$ implies that $m(A)=0$. Consider the set functions $m_n(A)=m(T^n A)$ for $n=1, 2, \ldots$ and for every measurable A. We see immediately that $m_n(A)$ is not necessarily additive and hence is not a measure on $(S, \mathfrak{B})$. Thus the procedure for defining the weight functions $w_n(x)$ cannot be followed here as previously and has to be modified. In this modification we are governed by the fact that our proposed theorem must reduce to the known special cases, i. e. to Theorem I in case T is one-to-one and to BIRKHOFF's theorem in case T is measure preserving in the sense that $m(T^{-1}A)=m(A)$. (Cf. F. RIESZ [9].) In fact let $\mathfrak{B}_1$ be the collection of all sets which are full inverse images of sets belonging to $\mathfrak{B}$. It is quite easy to see that $\mathfrak{B}_1$ is a Borel field of sets. It is also quite easy to see that $m_1(A)=m(TA)$ is a completely additive set function on $\mathfrak{B}_1$ and thus both m and m_1 are measures defined on $(S, \mathfrak{B}_1)$. Moreover m_1 is absolutely continuous with respect to m on $(S, \mathfrak{B}_1)$. Thus by using the RADON—NIKODYM theorem we see that there exists a $\mathfrak{B}_1$-measurable point function $w_1(x)$ such that

$$m_1(A)=\int_A w_1(x)\,dm,$$

for any $A \in \mathfrak{B}_1$. $w_1(x)$ is positive almost everywhere and without loss of generality we can assume that $w_1(x)$ is positive everywhere.

Let us now define $w_n(x)=w(T^{n-1}x)\ldots w(Tx)\,w(x)$, $n=2, 3, \ldots$. By considering approximating sums to the integrals in question one can see that

$$\int_S f(x)\,dm=\int_S f(Tx)\,w_1(x)\,dm$$

for every measurable function $f(x)$ such that either its positive or negative part is integrable. It follows that

$$\int_S f(x)\,dm=\int_S f(T^n x)\,w_n(x)\,dm \tag{1'}$$

for $n=0, 1, 2, \ldots$ and for every $f(x)$ which is described above. By definition we have

$$w_{i+j}(x)=w_i(T^j x)\,w_j(x) \quad \text{for} \quad i, j=0, 1, \ldots. \tag{2'}$$

With $w_n(x)$ as weight functions we form for every $\mathfrak{B}$-measurable function $q(x)$ the sum

$$q^n(x)=q(x)+q(Tx)\,w_1(x)+\ldots+q(T^{n-1}x)\,w_{n-1}(x).$$

Theorem II. *If T is a single valued measurable and non-singular point transformation of $(S, \mathfrak{B}, m)$ onto itself, if $f(x)$ is integrable and if $h(x)$ is non-negative and such that $h^n(x) \to \infty$ almost everywhere, then $f^n(x)/h^n(x)$ converges almost everywhere to a finite limit.*

The proof of Theorem II follows exactly the same lines as that of Theorem I and we shall therefore omit it here.

The question now arises as to when is it true that $h^n(x) \to \infty$ almost everywhere if, for instance, $h(x)$ is positive almost everywhere. In case T is one-to-one it was shown by HALMOS [2] (p. 157) that for $h(x) > 0$ almost everywhere, $h^n(x) \to \infty$ almost everywhere if there are no wandering sets of positive measure with respect to T, i. e. $T^i A \cap A = 0$ for $i = \pm 1, \pm 2, \ldots$ implies $m(A) = 0$. Thus the condition that $h^n(x) \to \infty$ can, at least for the case of a one-to-one transformation, be replaced by a condition which reflects directly on the nature of the transformation. In the more general case of a single valued transformation which is not one-to-one we have not been able to replace the condition $h^n(x) \to \infty$ by one directly bearing on the nature of T. We have been able to show that if there exists a measure μ on $(S, \mathfrak{B})$ which is invariant under T $(\mu(T^{-1}A) = \mu(A))$ and if T admits no wandering sets of positive measure then if $h(x) > 0$ almost everywhere, we have $h^n(x) \to \infty$. But in general the question of whether the condition that there are no wandering sets of positive measure under T (or some similar condition) yields $h^n(x) \to \infty$ for $h(x) > 0$ almost everywhere is still open.

Bibliography.

[1] BIRKHOFF, G. D., Proof of the ergodic theorem. *Proceedings National Academy of Sciences U. S. A.*, **17** (1931), pp. 656–660.

[2] HALMOS, P. R., An ergodic theorem, *Ibidem*, **32** (1946). pp. 156—161.

[3] HOPF, E., *Ergodentheorie* (Berlin, 1937), p. 49.

[4] HOPF, E., Über eine Ungleichung der Ergodentheorie, *Sitzungsberichte Math. Naturwiss. Abteilung der Bayerischen Akademie der Wiss.*, **1944** (1947), pp. 171—176.

[5] HUREWICZ, W., Ergodic theorems without invariant measure, *Annals of Math.*, **45** (1944), pp. 192—206.

[6] KHINTCHINE, A, Zur Birkhoffschen Lösung des Ergodenproblems, *Math. Annalen*, **107** (1932), pp. 385—488.

[7] OXTOBY, J. C., On the ergodic theorem of Hurewicz, *Annals of Math.*, **49** (1948), pp. 872—884.

[8] PITT, H. R., Some generalization of the ergodic theorem, *Proceedings Cambridge Philosophical Society*, **38** (1942), pp. 325—343.

[9] RIESZ, F., Sur quelques problèmes de la théorie ergodique, *Mat. Fiz. Lapok*, **49** (1942), pp. 34—62 (en hongrois); Sur la théorie ergodique, *Commentarii Math. Helvetici*, **17** (1945), pp. 221—239.

INSTITUTE FOR ADVANCED STUDY.

(Received July 6, 1949.)

Une généralisation du théorème tauberien de Wiener.

Par S. MANDELBROJT et S. AGMON à Paris.

1. Nous désignons par L la classe de toutes les fonctions $K(x)$, mesurables sur $(-\infty, \infty)$ et telles que $\int_{-\infty}^{\infty} |K(x)|\,dx < \infty$. Nous désignons par $g(u) = S(K)$ la transformée de Fourier de K:

$$g(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} K(x) e^{-iux}\,dx.$$

L'ensemble des racines de $g(u) = 0$ sera désigné par $\Omega(K)$. E étant un ensemble linéaire fermé quelconque, nous désignerons par E_f la frontière de E, c'est-à-dire l'ensemble de tous les points non intérieurs de E. Nous désignerons enfin par B la classe de toutes les fonctions mesurables et bornées sur $(-\infty, \infty)$.

2. Le théorème tauberien général de N. WIENER est bien connu:

Théorème W_1. *Si la relation*

$$\lim_{y=\infty} \int_{-\infty}^{\infty} K_0(y-x)h(x)\,dx = A \int_{-\infty}^{\infty} K_0(x)\,dx \tag{1}$$

a lieu pour une fonction $K_0 \in L$ avec $\Omega(K_0)$ vide, et une fonction $h \in B$, la relation

$$\lim_{y=\infty} \int_{-\infty}^{\infty} K(y-x)\,h(x)\,dx = A \int_{-\infty}^{\infty} K(x)\,dx \tag{2}$$

a lieu pour toute fonction $K \in L$.

Par contre si $\Omega(K_0)$ n'est pas vide il existe une fonction h de B et une fonction K de L telles que (1) ait lieu sans que (2) ait lieu. WIENER a démontré que le théorème W_1 résulte du théorème suivant:

Théorème W_2. *Soit $\{\xi_n\}$ une suite partout dense sur $(-\infty, \infty)$, et soit $K_0 \in L$ avec $\Omega(K_0)$ vide. Soit $K \in L$. A tout $\varepsilon > 0$ correspondent un entier N et* *des constantes* $a_1, a_2, \ldots a_N$ *tels que*

$$\int_{-\infty}^{\infty} \left| K(x) - \sum_{1}^{N} a_\nu K_0(x - \xi_\nu) \right| dx < \varepsilon. \tag{3}$$

Nous nous posons le problème général suivant: la fonction $K_0 \in L$ étant donnée, déterminer le sous-espace $\pi(K_0)$ de L, tel que pour toute fonction K de $\pi(K_0)$ une inégalité de la forme (3) ait lieu pour tout $\varepsilon > 0$. On voit immédiatement qu'une condition *nécessaire* pour que $K \in \pi(K_0)$ est que

$$\Omega(K) \supset \Omega(K_0). \tag{4}$$

Cette condition, est-elle aussi suffisante pour que $K \in \pi(K_0)$? Nous n'avons pu le démontrer qu'en ajoutant une hypothèse portant, soit sur $\Omega(K)$, soit sur la croissance de K (pour ne citer qu'un cas particulier d'un théorème plus général: la réponse est affirmative si $xK(x) \in L$)

Il résulte de la méthode générale de F. Riesz qu'une condition nécessaire et suffisante pour que (3) ait lieu pour tout $\varepsilon > 0$ est que le système infini d'équations

$$\int_{-\infty}^{\infty} K(x) \varphi_0(x)\, dx = 1, \quad \int_{-\infty}^{\infty} K_0(x - \xi_n) \varphi_0(x)\, dx = 0 \quad (n \geqq 1)$$

n'ait pas de solution $\varphi_0 \in B$.

La suite $\{\xi_n\}$ étant partout dense sur $(-\infty, \infty)$, on voit ainsi qu'une condition nécessaire et suffisante pour que $K \in \pi(K_0)$ est que toute fonction $\varphi_0 \in B$ vérifiant l'équation

$$\int_{-\infty}^{\infty} \varphi_0(x) K_0(y - x)\, dx = 0 \tag{5}$$

vérifie aussi l'équation

$$\int_{-\infty}^{\infty} \varphi_0(x) K(y - x)\, dx = 0. \tag{6}$$

3. A la fonction $\varphi \in B$ faisons correspondre, avec Carleman[1]), les fonctions suivantes

$$\begin{aligned} F^+(z) &= \int_{-\infty}^{0} \varphi(t) e^{-itz}\, dt, \quad I(z) > 0 \\ F^-(z) &= -\int_{0}^{\infty} \varphi(t) e^{-itz}\, dt, \quad I(z) < 0 \end{aligned} \tag{7}$$

respectivement holomorphes dans le demi-plan supérieur et demi-plan inférieur. D'après un théorème de Carleman, si $\varphi \in B$ est une solution de (6) $(K \in L)$, les fonctions $F^+(z)$ et $F^-(z)$, définies par (7) sont aussi régulières en tout

---

[1]) T. Carleman, *Intégrale de Fourier et questions qui s'y rattachent* (Uppsale, 1944).

point qui n'appartient pas à $\Omega(K)$, et chacune d'elles constitue le prolongement analytique de l'autre à travers tout segment ne contenant pas de points de $\Omega(K)$.

Nous allons maintenant démontrer le lemme suivant:

Lemme 1. *Soient $K \in L$, $\varphi_0 \in B$, et posons*

$$\varphi(x) = \int_{-\infty}^{\infty} \varphi_0(t) K(x-t)\,dt. \tag{8}$$

Si $\Omega(K_0)$ contient un intervalle (α, β), chacune des fonctions F^+, F^- définies par (7) *à partir de la fonction φ donnée par* (8) *(cette fonction est continue et $\varphi \in B$) est le prolongement analytique de l'autre à travers (α, β).*

La fonction continue $\Delta(x)$ définie par: $\Delta(x) = 1 - |x|$ pour $|x| \leq 1$, $\Delta(x) = 0$ pour $|x| \geq 1$, est la transformée de Fourier de $\delta(x)$ où $\delta(2x) = (2\pi)^{-\frac{1}{2}} (\sin x)^2/x^2$. La fonction $\Delta_0(x) = \Delta[(x-\xi)/l]$ où $\xi = (\alpha+\beta)/2$, $l = (\beta-\alpha)/2$, est par conséquent, la transformée de Fourier de $\delta_0(x) = l e^{i\xi x} \delta(lx)$.

Or on a

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi(t) \delta_0(x-t)\,dt &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \delta_0(x-t)\,dt \int_{\infty}^{\infty} \varphi_0(u) K(t-u)\,du = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi_0(u)\,du \int_{-\infty}^{\infty} \delta_0(x-t) K(t-u)\,dt = \int_{-\infty}^{\infty} \varphi_0(u) K^*(x-u)\,du \end{aligned} \tag{9}$$

où K^* est le produit symbolique ("faltung") de δ_0 et K. Comme $g^*(u) = S(K^*) = g(u)\Delta_0(u)$, où $g(u) = S(K)$, on voit que $g^*(u) \equiv 0$, donc $K^*(x) \equiv 0$, et (9) nous permet d'écrire

$$\int_{-\infty}^{\infty} \varphi(t) \delta_0(x-t)\,dt = 0.$$

Comme $\Delta_0(u)$ ne s'annule pas dans (α, β) le resultat énoncé résulte immédiatement du théorème cité de Carleman.

4. Nous pouvons maintenant démontrer le théorème suivant:

Théorème 1. *Soient $\varphi_0 \in B$, $K_0 \in L$, $K \in L$, φ_0 et K_0 étant liés par* (5). *Supposons que* (4) *ait lieu. La fonction φ définie par* (8) *satisfait alors aussi à l'équation* (8), *appartient à B et est continue; et la fonction définie par*

$$F(z) = \int_{-\infty}^{0} \varphi(t) e^{-itz}\,dt, \quad I(z) > 0, \tag{10}$$

est holomorphe dans tout le plan excepté un ensemble Σ qui est un sous-

ensemble parfait de l'ensemble

$$I(K_0, K) = \Omega_f(K_0) \cap \Omega_f(K).\text{[2]} \tag{11}$$

La fonction F peut aussi être définie par

$$F(z) = -\int_0^{\infty} \varphi(t) e^{-itz} dt, \quad I(z) < 0. \tag{12}$$

L'affirmation concernant φ est évidente. Il résulte alors du théorème de CARLEMAN cité et du lemme I que $\Sigma \subset I(K_0, K)$, la fonction F étant uniforme. Il nous reste à démontrer que l'ensemble Σ est parfait. Comme cet ensemble est fermé il faut démontrer qu'il ne contient pas de points isolés. Posons $M = \overline{\text{borne}} |\varphi(x)|$. Il résulte de (10) et (12) que pour $|y| > 0$:

$$|F(x+iy)| \leq \frac{M}{|y|}. \tag{13}$$

Supposons, contrairement à notre affirmation, qu'il existe un point isolé ξ de Σ. Il résulte alors de (13) que ξ est un pôle simple. Soit $l_0 > 0$ tel que l'intervalle $|x - \xi| \leq l_0$ ne contienne d'autres points de Σ que ξ, soit $0 < l < l_0$ et soit a le résidu de ξ. On a pour $\alpha > 0$:

$$\int_{\xi-l}^{\xi+l} \Delta\left(\frac{x-\xi}{l}\right) [F(x+i\alpha) - F(x-i\alpha)] dx = \int_{\xi-l}^{\xi+l} \Delta\left(\frac{x-\xi}{l}\right) dx \int_{-\infty}^{\infty} \varphi(t) e^{-\alpha|t|} e^{-itx} dt =$$

$$= \int_{-\infty}^{\infty} \varphi(t) e^{-\alpha|t|} dt \int_{\xi-l}^{\xi+l} \Delta\left(\frac{x-\xi}{l}\right) e^{-itx} dx = l\sqrt{2\pi} \int_{-\infty}^{\infty} e^{-\alpha|t|} \varphi(t) e^{-it\xi} \delta(lt) dt.$$

Comme $\varphi(t)\delta(lt) \in L$, on peut écrire

$$\lim_{\alpha=0} \int_{\xi-l}^{\xi+l} \Delta\left(\frac{x-\xi}{l}\right) [F(x+i\alpha) - F(x-i\alpha)] dx = l\sqrt{2\pi} \int_{\infty}^{\infty} \varphi(t) e^{-i\xi t} \delta(lt) dt. \tag{14}$$

Mais, en écrivant $F(z) = a(z-\xi)^{-1} + G(z)$, on sait que $G(z)$ est holomorphe pour $|x - \xi| \leq l$ On a donc

$$\begin{aligned} l\int_{-\infty}^{\infty} \varphi(t) e^{-i\xi t} \delta(lt) dt &= \lim_{\alpha=0} \frac{-1}{\sqrt{2\pi}} \int_{\xi-l}^{\xi+l} \Delta\left(\frac{x-\xi}{l}\right) \frac{2ai\alpha}{(x-\xi)^2 + \alpha^2} dx = \\ &= -\frac{2ai}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{dx}{1+x^2} = -\sqrt{2\pi} ai. \end{aligned} \tag{15}$$

En tenant compte de de (8) et du fait que $\xi \in \Omega(K)$, c'est-à-dire

$$\int_{-\infty}^{\infty} \delta(lu)\, e^{-i\xi t} K(t-u) dt = \delta(lu) e^{-i\xi u} g(\xi) = 0 \quad (g = S(K)),$$

[2]) C'est-à-dire que $I(K_0, K)$ est l'intersection des ensembles $\Omega_f(K_0)$ et $\Omega_f(K)$. Cet ensemble est donc fermé, partout non-dense.

on obtient:

$$(16)\quad l\int_{-\infty}^{\infty}\varphi(t)e^{-i\xi t}\delta(lt)\,dt = l\int_{-\infty}^{\infty}e^{-i\xi t}\delta(lt)\,dt\int_{-\infty}^{\infty}K(t-u)\varphi_0(u)\,du =$$

$$= l\int_{-\infty}^{\infty}\varphi_0(u)\,du\int_{-\infty}^{\infty}e^{-i\xi t}\delta(lt)K(t-u)\,dt = l\int_{-\infty}^{\infty}\varphi_0(u)\,du\int_{-\infty}^{\infty}e^{-i\xi t}[\delta(lt)-\delta(lu)]K(t-u)\,dt.$$

Il résulte alors de (15) et (16) que

$$\sqrt{2\pi}\,|a| \leq Ml\int_{-\infty}^{\infty}du\int_{-\infty}^{\infty}|\delta(lt)-\delta(lu)|\,|K(t-u)|\,dt =$$

$$= M\int_{-\infty}^{\infty}|K(t)|\,dt\int_{-\infty}^{\infty}|\delta(y+lt)-\delta(y)|\,dy,\quad 0<l<l_0.$$

En faisant tendre l vers zéro, on obtient $a=0$, ce qui prouve que ξ n'appartient pas à Σ et notre théorème est démontré.

Comme un ensemble parfait dénombrable est vide, il résulte du théorème I que si (4) a lieu et si $I(K_0, K)$ est dénombrable, $F(z)$ est une fonction entière, et d'après (13) $F(z)\equiv 0$, c'est-à-dire $\varphi(x)\equiv 0$. On a ainsi le théorème suivant:

Théorème II. *Si $K_0 \in L$, $K \in L$, si* (4) *a lieu et si l'ensemble* $I(K_0, K)$ *défini par* (11) *est dénombrable, la fonction K appartient à $\pi(K_0)$. Donc chaque solution φ_0 de* (5) *est aussi une solution de* (6)*;* (3) *est satisfait pour chaque $\varepsilon>0$; et si pour une fonction $h\in B$ on a* (1) *on a aussi* (2) *(généralisation du théorème tauberien de Wiener).*

5. Le théorème précédent peut être généralisé de la manière suivante:

Théorème III. *Les conclusions du théorème* II *subsistent si, en conservant les autres hypothèses de ce théorème, l'hypothèse que $I(K_0, K)$ est dénombrable est remplacée par la suivante: Il existe un ensemble ouvert O contenant $\Omega(K_0)$, sauf peut-être un sous-ensemble dénombrable de $\Omega(K_0)$, et possédant les propriétés suivantes:*

a) *à tout point $\xi\in O$ on peut faire correspondre un $\alpha>0$ tel que $g(u)=S(K)$ soit absolument continue sur $|u-\xi|\leq\alpha$.*

b) *on a*

$$(17)\qquad \int_{\xi-\alpha}^{\xi+\alpha}|g'(u)|^2\,du<\infty$$

[de a) il résulte que $g'(u)$ existe p. p. dans $|u-\xi|\leq\alpha$].

Les conditions a) et b) sont satisfaites si, par exemple, g est localement lipschitzienne d'ordre un, donc si $g'(u)$ est une fonction continue et, en particulier, si $xK(x)\in L$.

Pour la démonstration du théorème remarquons (nous conservons les notations de la démonstration du théorème II) que $F(z)$ est holomorphe en dehors de l'ensemble $\Omega(K_0)$ Nous démontrerons que cette fonction est aussi régulière en tout point de O, elle ne pourrait donc être singulière qu'en un ensemble dénombrable, et la conclusion du théorème III résultera, comme celle du théorème II, du théorème I.

Si donc ξ et α sont définis comme dans l'énoncé, définissons $\omega_\alpha(u)$ de la manière suivante: $\omega_\alpha(u)=1$ pour $|u-\xi|\leqq\alpha/2$, $\omega_\alpha(u)=0$ pour $|u-\xi|\geqq\alpha$, $\omega_\alpha(u)$ est linéaire pour $\alpha/2\leqq|u-\xi|\leqq\alpha$. $\omega_\alpha(u)$ est la transformée de Fourier d'une fonction $P_\alpha(x)$ appartenant à L. Posons $g_\alpha(u)=g(u)\omega_\alpha(u)$. $g_\alpha(u)$ est la transformée de Fourier de la fonction

$$K_\alpha(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}K(t)P_\alpha(x-t)dt \tag{18}$$

Posons

$$\varphi_\alpha(x)=\int_{-\infty}^{\infty}\varphi_0(t)K_\alpha(x-t)dt. \tag{19}$$

La fonction $g_\alpha(u)$ satisfait aussi aux conditions a) et b) et l'on a $g_\alpha(u)=0$ pour $|u-\xi|\geqq\alpha$. Désignons par Ω_α la partie de $\Omega(K_0)$ située dans $|u-\xi|\leqq\alpha$ et soit $m(\Omega_\alpha)$ sa mesure. Quel que soit $\varepsilon>0$, on peut couvrir Ω_α par un nombre fini, n, d'intervalles $[\alpha_i,\beta_i]$ $(i=1,\ldots,n)$:

$$\xi-\alpha\leqq\alpha_1<\beta_1<\alpha_2\ldots<\alpha_i<\beta_i<\ldots\leqq\xi+\alpha,$$

dont les extrémités appartiennent à Ω_α et tels que $\sum_1^n(\beta_i-\alpha_i)<m(\Omega_\alpha)+\varepsilon$. Posons $\beta_0=\xi-\alpha$, $\alpha_{n+1}=\xi+\alpha$ et soit $g_{\alpha,\varepsilon}(u)$ la fonction définie de la manière suivante: $g_{\alpha,\varepsilon}(u)=0$ pour $|u-\xi|\geqq\alpha$, $g_{\alpha,\varepsilon}(u)=0$ pour $\beta_i\leqq u\leqq\alpha_{i+1}$ $(i=0,\ldots,n)$, $g_{\alpha,\varepsilon}(u)=g_\alpha(u)$ pour $\alpha_i\leqq u\leqq\beta_i$ $(i=1,2,\ldots,n)$. La fonction $g_{\alpha,\varepsilon}(u)$ possède une dérivée p p., égale à $g'_\alpha(u)$ p. p. dans les intervalles (α_i,β_i) et égale à zéro ailleurs. Comme, d'autre part, $g'_\alpha(u)=0$ p. p. dans Ω_α, on voit que $g'_{\alpha,\varepsilon}(u)\neq0$ seulement sur un ensemble E_ε dont la mesure ne dépasse pas ε. D'ailleurs $g_{\alpha,\varepsilon}(u)$ ne diffère, elle-même, de zéro que sur un ensemble de mesure non supérieure à ε. Cette fonction est la transformée de Fourier de la fonction

$$K_{\alpha,\varepsilon}(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}g_{\alpha,\varepsilon}(u)e^{iux}du=\frac{1}{\sqrt{2\pi}}\int_{\xi-\alpha}^{\xi+\alpha}g_{\alpha,\varepsilon}(u)e^{iux}du. \tag{20}$$

En intégrant cette intégrale par parties on obtient (car $g_{\alpha,\varepsilon}(\xi\pm\alpha)=0$):

$$K_{\alpha,\varepsilon}(x)=\frac{1}{\sqrt{2\pi}x}\int_{\xi-\alpha}^{\xi+\alpha}g'_{\alpha,\varepsilon}(u)\,e^{iux}dx.$$

Et, en utilisant l'inégalité de Schwartz et l'égalité de Parseval, on a

$$\int_1^\infty |K_{\alpha,\varepsilon}(x)|\,dx \leqq \left\{\frac{1}{2\pi}\int_1^\infty \frac{dx}{x^2}\right\}^{1/2}\left\{\int_1^\infty\left|\int_{\xi-\alpha}^{\xi+\alpha} g'_{\alpha,\varepsilon}(u)e^{iux}du\right|^2 dx\right\}^{1/2}$$

$$\leqq \frac{1}{\sqrt{2\pi}}\left\{\int_{-\infty}^{\infty}\left|\int_{\xi-\alpha}^{\xi+\alpha} g'_{\alpha,\varepsilon}(u)\,e^{iux}\,du\right|^2 dx\right\}^{1/2} = \left\{\int_{\xi-\alpha}^{\xi+\alpha}|g'_{\alpha,\varepsilon}(u)|^2du\right\}^{1/2} = \left\{\int_{E\varepsilon}|g'_\alpha(u)|^2du\right\}^{1/2}.$$

La même inégalité est valable pour $\int_{-\infty}^{-1}|K_{\alpha,\varepsilon}(x)|dx$, et l'on a aussi $|K_{\alpha,\varepsilon}(x)|\leqq\varepsilon H$ où $H=(2\pi)^{-1/2}\underset{|u-\xi|\leqq\alpha}{Max}|g(u)|$. Il en résulte que

$$\lim_{\varepsilon\to 0}\int_{-\infty}^{\infty}|K_{\alpha,\varepsilon}(x)|dx=0. \tag{21}$$

Or il résulte du théorème II et de la manière même dont la fonction $g_{\alpha,\varepsilon}(u)$ a été définie que

$$\varphi_\alpha(x)=\int_{-\infty}^{\infty}\varphi_0(x-t)K_\alpha(t)dt=\int_{-\infty}^{\infty}\varphi_0(x-t)K_{\alpha,\varepsilon}(t)dt.$$

On voit donc d'après (21) que $\varphi_\alpha(x)\equiv 0$. Par conséquent, on peut écrire

$$F(z)=\int_{-\infty}^{0}(\varphi(t)-\varphi_\alpha(t))e^{-itz}dt$$

où

$$\varphi(x)-\varphi_\alpha(x)=\int_{-\infty}^{\infty}\varphi_0(t)[K(x-t)-K_\alpha(x-t)]dt.$$

Comme la transformée de Fourier de $K(x)-K_\alpha(x)$ est nulle pour $|u-\xi|<\frac{\alpha}{2}$, on voit d'après le lemme I que la fonction $F(z)$ est holomorphe dans cet intervalle, d'où le résultat cherché.

6. Démontrons maintenant les lemmes suivants:

L e m m e II. *Soit $K\in L$, et posons*

$$G(x)=\frac{1}{\sqrt{2\pi}}\int_{-1}^{1}(1-u^2)^2e^{iux}du. \tag{22}$$

On a $G\in L$, $xG(x)\in L$ et, en posant pour $N>0$:

$$K_N(x)=N\int_{-\infty}^{\infty}K(t)\,G[N(x-t)]\,dt; \tag{23}$$

on a pour $\varphi_1\in B$:

$$\lim_{N=\infty}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\varphi_1(x)K_N(x)dx=\int_{-\infty}^{\infty}\varphi_1(x)K(x)dx.$$

Lemme III. *Soit $P>0$. Il existe une constante C (ne dépendant que de P et K) telle que, quelle que soit la fonction positive mesurable $A(x)$ avec $A(x)\leqq|x|/T$, où $T>P$, $A(x)\leqq 1$ $(-\infty<x<\infty)$, on ait:*

$$(24)\qquad \int_{-\infty}^{\infty}|K_N(x)|\,A(x)\,dx\leqq\frac{C}{T}\int_0^T dx\left(\int_x^{\infty}|K(u)|\,du+\int_{-\infty}^{-x}|K(u)|\,du\right).$$

La démonstration du lemme II est immédiate; quant au lemme III, on a, en désignant le premier membre de (24) par B_n,

$$B_n\leqq N\int_{-\infty}^{\infty}|K(u)|\int_{-\infty}^{\infty}A(x)|G[N(x-u)]|\,dx\,du\leqq C_1\left(\int_{-\infty}^{-T}|K|\,du+\int_T^{\infty}|K|\,du\right)+$$

$$+\frac{C_2N}{T}\int_{-T}^{T}|K(u)|\int_{-\infty}^{\infty}|G[N(x-u)]x|\,dx\,du\leqq C_1\left(\int_{-\infty}^{-T}|K|\,du+\int_T^{\infty}|K|\,du\right)+$$

$$+\frac{C_3}{T}\int_{-T}^{T}|K(u)u|\,du+\frac{C_4}{T}\int_{-T}^{T}|K(u)|\,du\leqq C\left(\int_{-\infty}^{-T}+\int_T^{\infty}|K|\,du+\frac{1}{T}\int_{-T}^{T}|K(u)u|\,du\right).$$

Il suffit d'intégrer par parties la dernière intégrale dans le dernier membre pour avoir le résultat cherché.

7. E étant un ensemble situé sur la droite et I étant un intervalle, partageons I en n parties égales, et désignons par $\mu_n(E,I)$ la longueur totale de ceux de ces intervalles qui contiennent des points de $E\cap I$. Quel que soit $\alpha<1$, il existe un ensemble parfait E tel que pour tout I, $\varliminf_{n=\infty}\mu_n(E,I)n^{\alpha}=0$. Ainsi l'ensemble composé de tous les points $x=\sum_{\nu=1}^{\infty}c_\nu/q^\nu$ où $q\geqq 3$ est un entier, et où les c_ν prennent de toutes les manières possibles, k valeurs entières positives: $\alpha_1<\alpha_2\ldots<\alpha_k<q$, est un ensemble parfait avec $\mu_{q^m}(E,I)=$ $=\left(\frac{k}{q}\right)^m=(q^m)^{\frac{\log k}{\log q}-1}$ pour tout intervalle I contenant $[0,1]$; autrement dit on a $\varliminf\mu_n(E,I)n^{\alpha}=0$ pour tout $\alpha<1-\log k/\log q$. L'exemple de Cantor correspond au cas $q=3$, $k=2$.

Théorème IV. *Soient $K_0\in L$, $K\in L$. Supposons que la condition (4) est satisfaite. Désignons par I_N l'intervalle $|x|\leqq N$, et posons $\mu_n^{(N)}=\mu_n(I(K_0,K),I_N)$. Si pour chaque $N>0$ on a*

$$(25)\qquad \varliminf_{n=\infty}\mu_n^{(N)}\int_0^n dx\left(\int_x^{\infty}|K(u)|\,du+\int_{-\infty}^{-x}|K(u)|\,du\right)=0$$

les conclusions du théorème II *subsistent.*

Ainsi, si $\varliminf_{n=\infty}\mu_n^{(N)}n^{\alpha}=0$, il suffit que $\int_{-\infty}^{-x}+\int_x^{\infty}|K(u)|du=O(x^{\beta})$ $(0<x\to\infty)$ avec $\beta\leqq\alpha-1$ pour que (25) soit satisfait.

Soit, pour $N>0$, K_N la fonction définie par (23) et posons $E_N=I(K_0, K)\cap I_N$. On a $E_N=I(K_0, K_N)$. Partageons I_N en n parties égales, désignons par $J_n^{(N)}$ l'ensemble de ces intervalles partiels qui contiennent des points de E_N, et soient $u_1<u_2<\ldots u_q$ les abscisses des extrémités et des milieux des intervalles de $J_n^{(N)}$. Soit p_n le nombre d'intervalles de $J_n^{(N)}$. Il est clair que $q\leqq 3p_n$.

Posons

$$\Delta_{n,N}(u)=\sum_{j=1}^{q}\Delta\left[\frac{n}{N}(u-u_j)\right],$$

la fonction Δ étant définie comme dans le n° 3. Il est clair que $\Delta_{n,N}(u)=1$ pour u appartenant à un intervalle de $J_n^{(N)}$, $\Delta_{n,N}(u)=0$ lorsque u est à une distance non inférieure à N/n d'un tel intervalle, et $\Delta_{n,N}(u)$ est une fonction linéaire ailleurs. Soit $g_N(u)$ la transformée de Fourier de K_N, et posons

$$G_{n,N}(u)=g_N(u)\Delta_{n,N}(u)=\sum_{j=1}^{q}\left\{\Delta\left[\frac{n}{N}(u-u_j)\right]\left[g_N(u)-g_N(u_j)\right]+\right.$$
$$\left.+\Delta\left[\frac{n}{N}(u-u_j)\right]g_N(u_j)\right\}.$$

Or le premier membre de l'accolade est la transformée de Fourier de

$$Q_{N,n,j}(x)=\frac{N}{n\sqrt{2\pi}}\int_{-\infty}^{\infty}K_N(x-y)\,e^{iu_jy}\,\delta\left(\frac{Ny}{n}\right)dy-\frac{N}{n}g_N(u_j)\,e^{iu_jx}\,\delta\left(\frac{Nx}{n}\right)=$$

(26)

$$=\frac{N}{n\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{iu_j(x-y)}\left[\delta\left(\frac{N}{n}(x-y)\right)-\delta\left(\frac{N}{n}x\right)\right]K_N(y)\,dy,$$

et le second membre de cette accolade est la transformée de Fourier de

(27) $$R_{N,n,j}(x)=\frac{N}{n}g_N(u_j)\,e^{iu_jx}\,\delta\left(\frac{N}{n}x\right).$$

Mais, d'après la définition de u_j, il existe un point u_j' tel que $|u_j'-u_j|\leqq\frac{2N}{n}$ avec $g_N(u_j')=0$. Par conséquent:

$$|g_N(u_j)|=|g_N(u_j)-g_N(u_j')|\leqq\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}|K_N(x)(e^{-iu_jx}-e^{-iu_j'x})|dx$$

Comme $|e^{-iu_jx}-e^{-iu_j'x}|\leqq|(u_j-u_j')x|\leqq\frac{2N}{n}|x|$ et $|e^{-iu_jx}-e^{-iu_j'x}|\leqq 2$ on voit, d'après le lemme III, avec $A(x)=|e^{-iu_jx}-e^{-iu_j'x}|/2(N+1)$, $T=n$ que $|g_N(u_j)|\leqq CB(n)$, où l'on a posé

$$B(T)=\frac{1}{T}\int_0^T dx\left(\int_x^{\infty}|K(u)|du+\int_{-\infty}^{-x}|K(u)|du\right).$$

Il résulte donc de (27) que

$$\int_{-\infty}^{\infty} |R_{N,n,j}(x)|\,dx \leq C_2 B(n). \tag{28}$$

Il résulte aussi de (26) et du lemme III avec $A(x) = \alpha \int_{-\infty}^{\infty} \left|\delta\left(y - \frac{N}{n}x\right) - \delta(y)\right| dy$, où α est une constante convenablement choisie, que

$$\begin{aligned}\int_{-\infty}^{\infty} |Q_{N,n,j}(x)|\,dx &\leq \frac{N}{n\sqrt{2\pi}} \int_{-\infty}^{\infty} |K_N(x)| \int_{-\infty}^{\infty} \left|\delta\left[\frac{N}{n}(y-x)\right] - \delta\left(\frac{N}{n}y\right)\right| dy\,dx = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |K_N(x)| \int_{-\infty}^{\infty} \left|\delta\left(y - \frac{N}{n}x\right) - \delta(y)\right| dy\,dx \leq C_3 B(n).\end{aligned} \tag{29}$$

Ainsi $G_{N,n}(u)$ est la transformée de Fourier d'une fonction $K_{N,n}(x)$ telle que

$$\int_{-\infty}^{\infty} |K_{N,n}(x)|\,dx \leq C_4 q_n B(n) \leq C_5 p_n B(n) = C_5 \mu_n^{(\lambda)} \int_0^n dx \left(\int_{-\infty}^{-x} |K|\,du + \int_x^{\infty} |K|\,du\right).$$

Il résulte donc de nos hypothèses que

$$\lim_{n=\infty} \int_{-\infty}^{\infty} |K_{N,n}(x)|\,dx = 0. \tag{30}$$

La fonction $G^*_{N,n}(u) = g_N(u) - G_{N,n}(u)$ (où $g_N(u) = S(K_N)$) s'annule dans un ensemble d'intervalles qui, dans leur réunion, contiennent $I(K_0, K_N)$. Si $K^*_{N,n}(x)$ désigne la fonction dont $G^*_{N,n}$ est la transformée de Fourier, l'ensemble $I(K_0, K^*_{N,n})$ ne contient que des points isolés, et d'après le théorème II

$$\int_{-\infty}^{\infty} K^*_{N,n}(y-x)\varphi_0(x)\,dx = 0.$$

La relation (30) fournit alors

$$\int_{-\infty}^{\infty} K_N(y-x)\varphi_0(x)\,dx = 0,$$

et le lemme II permet alors d'écrire

$$\int_{-\infty}^{\infty} K(y-x)\varphi_0(x)\,dx = 0.$$

(Reçu le 24 septembre 1949)

Valeurs propres et vecteurs propres d'un endomorphisme complètement continu d'un espace vectoriel à voisinages convexes.

Par JEAN LERAY à Paris.

Dans son célèbre Mémoire [7], M. FRÉDÉRIC RIESZ a étendu la théorie de FREDHOLM aux équations linéaires des espaces de Banach; cette extension fut si utile qu'on désire l'adapter aux tendances actuelles: substituer les espaces topologiques aux espaces métriques et en particulier les espaces linéaires à voisinages convexes aux espaces de Banach. C'est cette adaptation que nous allons effectuer. L' "alternative de Fredholm" résultera d'un théorème de topologie, dont la preuve utilise essentiellement des applications non linéaires: le théorème de l'invariance du domaine; BROUWER l'établit pour les espaces euclidiens; il a été étendu aux espaces abstraits par le profond mathématicien polonais J. SCHAUDER, qui fut victime des massacres nazis; je l'ai récemment prouvé pour un type d'espaces englobant les espaces à voisinages convexes. Les autres théorèmes seront obtenus par les raisonnements-mêmes de F. RIESZ, tels que la connaissance de l'alternative de Fredholm permet de les simplifier; mais nous devrons justifier ces raisonnements par des lemmes autres que ceux de F. RIESZ.

I. ESPACE VECTORIEL À VOISINAGES CONVEXES.

1. Définitions. Par *espace vectoriel* nous entendrons espace vectoriel sur le corps A des nombres réels ou complexes, au sens de [1]. Soit X un tel espace. Le segment joignant les points x et y de X est l'ensemble des points $\alpha x+(1-\alpha)y$ ou $0\leqq\alpha\leqq 1$ (α réel); une partie de X est dite convexe quand elle contient le segment joignant deux quelconques de ses points. Un *espace vectoriel à voisinages convexes* est un espace vectoriel X muni d'une topologie possédant les trois propriétés suivantes: *a)* cette topologie est *séparée* au sens de [2] (elle est même régulière, car X est uniforme); *b*) les applications $x+y$ et αx ($\alpha\in A$) sont *continues* par rapport à leurs arguments; *c)* le point 0 (et par suite tout point de X) possède *un système fondamental de voisinages convexes.*

Les *espaces euclidiens*, c'est-à-dire les produits d'un nombre fini de droites (réelles ou complexes selon que A est le corps des nombres réels ou complexes), constituent un exemple d'espaces vectoriels à voisinages convexes.

Notations. A est le corps des nombres réels ou complexes; X est un espace vectoriel à voisinages convexes. Soient B une partie de A; Y et Z des parties de X; BY est l'ensemble des points βy tels que $\beta \in B$, $y \in Y$; $Y \pm Z$ est l'ensemble des points $y \pm z$ tels que $y \in Y$, $z \in Z$. $\overline{Y}$ et $\dot{Y}$ désigneront l'adhérence et la frontière de Y. La négation de $\in$ est désignée par $\overline{\in}$.

2. Parties compactes. Lemme 2. 1. *Soient un voisinage V du point* 0, *une partie compacte K de X et une suite $x_1, x_2, \ldots$ de points de K tels que $x_\nu \overline{\in} x_\mu + V$ si $\mu < \nu$; cette suite est finie.*

Preuve. Sinon la réunion des points $x_1, x_2, \ldots$ serait une partie fermée, non compacte de K, contrairement à [2], ch. I, § 10, proposition 4.

Lemme 2. 2. *Soient un espace topologique T, un espace compact K, une application continue $\varphi(t, k)$ de $T \times K$ dans X et une partie fermée F de X étrangère à $\varphi(t, K)$; le point t a un voisinage V tel que F soit étranger*[1]) *à $\varphi(V, K)$.*

Preuve. Soit $k \in K$; il existe des voisinages $V(k)$ de t et $W(k)$ de k tels que F soit étranger à $\varphi(V(k), W(k))$; on recouvre K par un nombre fini de $W(k)$; V est l'intersection des $V(k)$ correspondants.

Lemme 2. 3. *Étant donnés une partie compacte K de X et un voisinage V du point* 0, *il existe un voisinage Δ du nombre* 0 *tel que $\Delta K \subset V$.*

Preuve. On applique le lemme 2. 2 à $\varphi = \alpha k$.

Lemme 2. 4. *Si F est une partie fermée et K une partie compacte de X, alors $F + K$ est fermé.*

Preuve. Soit $x \overline{\in} F + K$; F est étranger à $x - K$; d'après le lemme 2. 2, x a un voisinage V tel que F soit étranger à $V - K$; V est donc étranger à $F + K$.

3. Parties fermées et ouvertes. Lemme 3. 1. *Si B est un ensemble fermé de nombres $\neq 0$ et si F est un ensemble fermé de points de $X \neq 0$, alors BF est fermé.*

Preuve. Soit Γ la partie compacte de A que constituent le nombre 0 et les nombres β^{-1} tels que $\beta \in B$; soit $x \overline{\in} BF$; F est étranger à Γx; d'après le lemme 2. 2, x possède un voisinage V tel que F soit étranger à ΓV; V est donc étranger à BF.

Lemme 3. 2. *Si W est un voisinage du point* 0 *et Δ un voisinage du nombre* 0, *l'intersection des βW tels que $\beta \overline{\in} \Delta$ est un voisinage V du point* 0.

1) "étranger à" signifie "sans point commun avec".

Preuve. On applique le lemme 3.1 aux complémentaires B et F de $\varDelta$ et W, supposés ouverts.

Le lemme suivant et évident:

Lemme 3.3. *Si P est une partie quelconque de X et G une partie ouverte de X, alors $P+G$ est ouvert.*

4. Sous-espaces. Par *sous-espace* de X nous entendons sous-espace vectoriel: tout sous-espace de X contient le point 0 et est espace vectoriel à voisinages convexes.

Lemme 4.1. (Adaptation de [7], Hilfssatz 1, 2, 3). *Supposons que le point 0 de X possède un voisinage ouvert V tel que $\overline{V}$ soit compact; soit Y un sous-espace fermé de X, différent de X; il existe $x \in \overline{V}$ tel que $x \overline{\in} Y+V$.*

Remarque. D'après le lemme 4.3, *X est euclidien; donc tous ses sous-espaces sont fermés.*

Preuve. Soit $z \in X$ tel que $z \overline{\in} Y$; puisque Y est fermé, il existe un voisinage W du point 0 tel que $z \overline{\in} Y+W$; le lemme 2.3 donne un nombre $\alpha \neq 0$ tel que $\alpha \overline{V} \subset W$; donc $z \overline{\in} Y+\alpha V$; $\alpha^{-1} z \overline{\in} Y+V$;

$$X \neq Y+V. \tag{4.1}$$

Supposons $\overline{V} \subset Y+V$; on aurait: $Y+\overline{V}=Y+V$; $Y+\overline{V}$ est fermé d'après le lemme 2.4; $Y+V$ est ouvert d'après le lemme 3.3; X est connexe; donc, contrairement à (4.1), $X=Y+V$.

Lemme 4.2. *Soit Y un sous-espace de X ayant une dimension finie, au sens de la théorie des espaces vectoriels:* [1], ch. II, § 3, n° 2; a) *Y est euclidien;* b) *Y est un sous-espace fermé de X.*

Preuve de a). Soit $(y_1, y_2, \ldots, y_\omega)$ une base de Y; soit E un espace euclidien de dimension ω; en associant au point de E de coordonnées $(\eta_1, \eta_2, \ldots \eta_\omega)$ le point $\eta_1 y_1 + \ldots \eta_\omega y_\omega$ de Y; on définit une application φ linéaire, continue et biunivoque de E sur Y; il s'agit de prouver qu'elle est bicontinue, c'est-à-dire qu'elle applique toute partie ouverte de E sur une partie ouverte de Y. Soit U la boule $|\eta_1|^2 + \ldots |\eta_\omega|^2 < 1$; il suffit de prouver que $\varphi(U)$ est un voisinage du point 0 de Y. Or $\dot{U}$ est compact et $0 \overline{\in} \dot{U}$; donc $\varphi(\dot{U})$ est compact ([2], ch. I, § 10, th. 1) et $0 \overline{\in} \varphi(\dot{U})$; le point 0 de Y a donc un voisinage V convexe et étranger à $\varphi(\dot{U})$; $\overset{-1}{\varphi}(V)$ est convexe, contient le point 0 et est étranger à la sphère $\dot{U}$, dont le centre est ce point 0; donc $\overset{-1}{\varphi}(V) \subset U$; donc $V \subset \varphi(U)$: $\varphi(U)$ est bien un voisinage du point 0 de Y.

Preuve de b). Supposons qu'il existe $z \in \overline{Y}$ tel que $z \overline{\in} Y$; soit Z le sous-espace de X ayant pour base $(z, y_1, \ldots y_\omega)$; Y serait un sous-espace non fermé de Z, qui serait euclidien d'après a); c'est impossible.

Lemme 4.3. *Tout sous-espace localement compact Y de X est fermé et euclidien.* (Voir un théorème plus général: [10] § 29.)

Nota. Y est localement compact quand le point 0 a, dans Y, un voisinage compact.

Preuve. Soit V un voisinage ouvert de 0 dans Y tel que $\overline{V}$ soit compact. Définissons par récurrence une suite de sous-espaces fermés et euclidiens Y_ν de Y: $Y_0=0$; si $Y_{\nu-1}\neq Y$, le lemme 4.1 donne un point $y_\nu \in \overline{V}$ tel que $y_\nu \bar{\in} Y_{\nu-1}+V$; Y_ν sera l'espace vectoriel de dimension ν contenant $Y_{\nu-1}$ et y_ν; d'après le lemme 4.2, Y_ν est un sous-espace fermé et euclidien de Y. On a $y_\nu \in \overline{V}$ et, si $\mu<\nu$, $y_\nu \bar{\in} y_\mu+V$; vu le lemme 2.1, la suite des Y_ν est donc finie: son dernier terme est Y; Y est donc sous-espace euclidien de X; Y est un sous-espace fermé de X d'après le lemme 4.2 b.

II. Endomorphisme de X complètement continu.

Adaptons comme suit la définition, due à F. Riesz, des applications complètement continues ([7] "vollstetig"):

5. Définition. *Soit $v(x)$ un endomorphisme de X, c'est-à-dire une application linéaire de X en lui-même; $v(x)$ est dite complètement continue si elle est continue et s'il existe un voisinage que $v(x)$ applique dans un compact: on a*

(5.1) $$v(\overline{W})\subset K,$$

W étant un voisinage ouvert du point 0 et K un compact.

Notations. Étant donné un endomorphisme de X complètement continu, $v(x)$, nous étudierons l'endomorphisme, dépendant du paramètre réel ou complexe $\lambda\in A$,

(5.2) $$u(x)=\lambda x-v(x).$$

Nous noterons parfois cet endomorphisme: $u=\lambda-v$. Nous poserons $u^2(x)=u(u(x)),\ldots, u^{\nu+1}(x)=u(u^\nu(x)),\ldots$; $\overset{-\nu}{u}(y)$ sera l'ensemble des points x tels que $u^\nu(x)=y$.

Nos conclusions permettront de discuter, quand $\lambda\neq 0$, *l'équation* d'inconnue $x\in X$, de paramètres $\lambda\in A$, $y\in X$: $u(x)=y$.

Lemme 5.1. *Si Y est un sous-espace fermé de X que v applique en lui-même, la restriction de v à Y est complètement continue.*

Preuve: $v(Y\cap\overline{W})\subset Y\cap K$; $Y\cap K$ est compact.

6. $u(X)$ est fermé. Lemme 6.1. *Si F est une partie fermée de $\overline{W}$, $u(F)$ est fermé.* (Voir un théorème plus général: [5] n⁰ 91, théorème 29 bis.)

Preuve. Soit $y\bar{\in} u(F)$; il s'agit de construire un voisinage de y étranger à $u(F)$. Soit V un voisinage fermé de y, étranger au compact $u(F\cap\lambda^{-1}(y+K))$; soit $F_1=F\cap \overset{-1}{u}(V)$; il suffit de construire un voisinage V_1 de y étranger à $u(F_1)$. Or F_1 est fermé et étranger à $\lambda^{-1}(y+K)$; $v(F_1)\subset K$ d'après (5.1); donc

$y \overline{\in} \lambda F_1 - K$ et, vu (5. 2), $u(F_1) \subset \lambda F_1 - K$; d'après le lemme 2. 4, $\lambda F_1 - K$ est fermé; V_1 sera son complémentaire.

Lemme 6. 2. ([7], Satz 5) *$u(X)$ est un sous-espace fermé de X.*

Preuve. $F_1 = \overset{-1}{u} u(\overline{W}) = \overline{W} + \overset{-1}{u}(0)$ est fermé, vu le lemme 6. 1; $O_1 = \overset{-1}{u} u(W) = W + \overset{-1}{u}(0)$ est ouvert, vu le lemme 3. 3. Soit F_2 le complémentaire de O_1 dans F_1; F_2 est fermé; $\dot{F}_1 \subset F_2$. Soit x un point étranger à F_1; le point 0 est intérieur à F_1; le segment joignant 0 à x, étant connexe, contient au moins un point de $\dot{F}_1$: en notant B l'ensemble des nombres réels $\beta \geqq 1$, nous avons $x \in B\dot{F}_1 \subset BF_2$; d'où $X = F_1 \cup BF_2$ et par suite

$$u(X) = u(F_1) \cup Bu(F_2) \tag{6. 1}$$

$u(F_1) = u(\overline{W})$ est fermé d'après le lemme 6. 1; $u(F_2)$ est le complémentaire de $u(W)$ dans $u(\overline{W})$; donc $u(F_2) = u(F_3)$, F_3 étant le complémentaire de $\overline{W} \cap O_1$ dans $\overline{W}$; $u(F_3)$ est fermé d'après le lemme 6. 1; $u(F_2)$ est donc fermé; $0 \overline{\in} u(F_2)$ car F_2 est étranger à $\overset{-1}{u}(0)$; donc, vu le lemme 3. 1, $Bu(F_2)$ est fermé. De (6. 1) résulte donc que $u(X)$ est fermé.

III. Valeurs propres.

7. L'alternative de Fredholm. Définition. Les *valeurs propres* de $v(x)$ sont les nombres λ tels que l'équation $u(x) = 0$, c'est-à-dire $\lambda x = v(x)$ possède des solutions $x \neq 0$.

Théorème 7. 1. ([7], Satz 7) *$u(x)$ est une application bicontinue de X sur lui-même, quand λ diffère de 0 et des valeurs propres de $v(x)$.*

Ce théorème résulte de l'extension du théorème de l'invariance du domaine aux espaces convexoïdes ([5], n° 95 et 96; théorème 36 et corollaire 36) ou, plus directement, aux espaces à voisinages convexes: [6]

Remarque. Le théorème 7. 1 implique *l'alternative de Fredholm:* Supposons $\lambda \neq 0$; ou bien l'équation $u(x) = 0$ possède une solution $x \neq 0$; ou bien l'équation $u(x) = y$ possède une solution x unique, quel que soit $y \in X$.

8. L'ensemble des valeurs propres. Théorème 8 1. ([7], Satz 12) *Les valeurs propres de $v(x)$ sont en nombre fini ou constituent une suite ayant pour limite 0.*

Preuve (empruntée à [7]). Soit une suite $\lambda_1, \ldots, \lambda_\nu, \ldots$ de valeurs propres distinctes, étrangères à un voisinage Δ du point 0; soit $x_\nu \neq 0$ tel que

$$\lambda_\nu x_\nu = v(x_\nu). \tag{8. 1}$$

Si x_ν dépendait linéairement de $x_1, x_2, \ldots x_{\nu-1}$, on aurait

$$\alpha_1 x_1 + \ldots \alpha_{\nu-1} x_{\nu-1} = x_\nu;$$

d'où, vu (8. 1), en transformant les deux membres par $\lambda_\nu - v$:

$$\alpha_1(\lambda_\nu - \lambda_1) x_1 + \ldots \alpha_{\nu-1}(\lambda_\nu - \lambda_{\nu-1}) x_{\nu-1} = 0;$$

il existerait donc $\mu < \nu$ tel que x_μ dépende linéairement de $x_1, \dots x_{\mu-1}$; d'où, par récurrence, $x_1 = 0$; or $x_1 \neq 0$; donc $x_1, x_2, \dots x_\nu$ sont linéairement indépendants. Soient X_ν le sous-espace qu'ils engendrent; X_ν est fermé et euclidien (lemme 4. 2);

$$X_\nu \neq X_{\nu+1}; \ X_1 \subset X_2 \subset \dots \subset X_\nu \subset X_{\nu+1} \subset \dots;$$

d'après le lemme 4. 1 il existe y_ν tel que

$$(8.2) \qquad y_\nu \in X_\nu \cap \overline{W}, \qquad (8.3) \qquad y_\nu \overline{\in} X_{\nu-1} + W;$$

de (8. 1) et de la définition de X_ν résulte, puisque $y_\nu \in X_\nu$,

$$(8.4) \qquad \lambda_\nu y_\nu - v(y_\nu) \in X_{\nu-1}; \ v(y_\mu) \in X_{\nu-1} \quad \text{pour } \mu < \nu;$$

de (8. 3) et (8. 4) résulte

$$\lambda_\nu y_\nu \overline{\in} \lambda_\nu y_\nu - v(y_\nu) + v(y_\mu) + \lambda_\nu W$$

c'est-à-dire $v(y_\nu) \overline{\in} v(y_\mu) + \lambda_\nu W$. D'après le lemme 3. 2, le point 0 a un voisinage V tel que $V \subset \lambda_\nu W$; donc $v(y_\nu) \overline{\in} v(y_\mu) + V$.

D'autre part, il résulte de (8. 2) et (5. 1) que $v(y_\nu) \in K$. Donc, en vertu du lemme 2. 1, la suite λ_ν doit être finie.

IV. VECTEURS PROPRES ET VECTEURS PRINCIPAUX.

Donnons à λ une valeur propre de v autre que 0.

9. Les sous-espaces $\overset{-\nu}{u}(0)$ et $u^\nu(X)$. Lemme 9. 1. ([7], Satz 1) $\overset{-1}{u}(0)$ *est un sous-espace fermé et euclidien;* $\overset{-1}{u}(0) \cap \overline{W}$ *est dans ce sous espace un voisinage compact du point* 0.

Preuve. $Y = \overset{-1}{u}(0)$ est fermé car $u(x)$ est continu. D'après (5. 2), $x = \lambda^{-1} v(x)$ sur Y; d'où $Y \cap \overline{W} = Y \cap \lambda^{-1} v(W)$; vu (5. 1), $Y \cap \overline{W}$ est donc un voisinage compact du point 0 de Y; d'après le lemme 4. 3, Y est donc euclidien.

Lemme 9. 2. ([7], Satz 1') $\overset{-\nu}{u}(0)$ *est est un sous-espace fermé et euclidien;* $\overset{-\nu}{u}(0) \cap \overline{W}$ *est dans ce sous-espace un voisinage compact du point* 0.

Preuve. On peut dans le lemme 9. 1 remplacer u par u^ν car $\lambda^\nu - u^\nu$, étant un polynome en v sans terme constant, est complètement continu.

Lemme 9. 3. ([7], Satz 2) *Il existe un entier σ tel que*

$$\overset{-1}{u}(0) \subset \overset{-2}{u}(0) \subset \dots \subset \overset{-\sigma}{u}(0) = \dots = \overset{-\tau}{u}(0) = \dots \quad (\sigma \leqq \tau).$$

Preuve (empruntée à [7]). Soit $\nu > 1$; on a $\overset{-\nu+1}{u}(0) \subset \overset{-\nu}{u}(0)$ car $u^{\nu-1}(x) = 0$ entraîne $u^\nu(x) = 0$. Supposons

$$(9.1) \qquad \overset{-\nu+1}{u}(0) \neq \overset{-\nu}{u}(0);$$

il existe d'après les lemmes 4. 1 et 9. 2 un point x_ν tel que

$$(9.2) \quad x_\nu \in \overset{-\nu}{u}(0) \cap \overline{W}; \qquad (9.3) \quad x_\nu \overline{\in} \overset{-\nu+1}{u}(0) + W.$$

D'après (9. 2), $u^\nu(x_\nu)=0$; donc

(9. 4) $$u(x_\nu)\in \overset{-\nu+1}{u}(0).$$

Pour $\mu<\nu$ il résulte de $u^\mu(x_\mu)=0$ que

(9. 5) $$x_\mu\in \overset{-\nu+1}{u}(0),\quad u(x_\mu)\in \overset{-\nu+1}{u}(0).$$

De (9. 3), (9. 4) et (9. 5) résulte

$$\lambda x_\nu \overline{\in} u(x_\nu)+\lambda x_\mu - u(x_\mu)+\lambda W,$$

c'est-à-dire, vu (5. 2), $v(x_\nu)\overline{\in} v(x_\mu)+\lambda W$.

D'autre part, d'après (9. 2) et (5. 1), $v(x_\nu)\in K$.

En vertu du lemme 2. 1, la suite des ν vérifiant (9. 1) est donc finie; σ est le dernier terme de cette suite.

Lemme 9. 4 $\overset{-\nu}{u}(0)\cap u^\sigma(X)=0$, *quel que soit* $\nu>0$.

Preuve (empruntée à [7]): Soit $x\in\overset{-\nu}{u}(0)\cap u^\sigma(X)$: il existe $y\in X$ tel que $x=u^\sigma(y)$; $u^\nu(x)=0$ implique $u^{\nu+\sigma}(y)=0$, c'est-à-dire $y\in \overset{\nu-\sigma}{u}(0)=\overset{-\sigma}{u}(0)$ (lemme 9. 3); d'où $u^\sigma(y)=0$, c'est-à-dire $x=0$.

Lemme 9. 5. ([7], Satz 6, 11) $u^\nu(X)$ *est un sous-espace fermé*;

$$X\supset u(X)\supset u^2(X)\supset\ldots\supset u^\sigma(X)=\ldots=u^\tau(X)=\ldots\qquad(\sigma\leqq\tau);$$

u est une application bicontinue de $u^\sigma(X)$ sur lui-même.

Preuve. $u(X)$ est fermé (lemme 6. 2). Soit $\nu>0$; supposons prouvé que $u^\nu(X)$ est fermé; le restriction de v à $u^\nu(X)$ applique $u^\nu(X)$ en lui-même car $vu^\nu=u^\nu v$; elle est complètement continue (lemme 5. 1); $u^{\nu+1}(X)$ est donc fermé (lemme 6. 2). La restriction de v à $u^\sigma(X)$ n'a pas λ pour valeur propre (lemme 9. 4 où $\nu=1$); donc (théorème 7. 1) la restriction de u à $u^\sigma(X)$ est une application bicontinue de $u^\sigma(X)$ sur lui-même; en particulier $u^{\sigma+1}(X)=u^\sigma(X)$.

Lemme 9. 6 ([7], Satz 8). $X=\overset{-\sigma}{u}(0)+u^\sigma(X)$, *cette somme étant directe.*

Preuve (empruntée à [7]) Soit $x\in X$: d'après le lemme 9. 5 il existe $y\in X$ tel que $u^{2\sigma}(y)=u^\sigma(x)$; d'où $u^\sigma(x-u^\sigma(y))=0$, c'est-à-dire $x\in\overset{-\sigma}{u}(0)+u^\sigma(X)$. Cette somme est directe d'après le lemme 9. 4, où l'on fait $\nu=\sigma$.

10. Vecteurs propres et principaux. Soit λ une valeur propre de v; les points x de X tels que $u(x)=0$ sont nommés *vecteurs propres de v correspondant à la valeur propre λ*; les points x de X auxquels on peut associer un entier $\nu>0$ tel que $u^\nu(x)=0$ ont été nommés par GOURSAT *vecteurs principaux de v correspondant à la valeur propre λ.*

Soit $v(x)$ un endomorphisme d'un espace euclidien E; le déterminant de la matrice de l'endomorphisme $\varrho x-v(x)$ $(\varrho\in A)$, par rapport à une base de E, est un polynome $\chi(\varrho)$, indépendant du choix de cette base; on nomme $\chi(\varrho)$ *fonction caractéristique* de v. On a ([9] chapitre XV, § 112):

(10. 1) $$\chi(v)=0.$$

Lemme 10. 1. *Soit u un endomorphisme nilpotent*[2]) *d'un espace euclidien E de dimension δ; sa fonction caractéristique est $\chi(\varrho)=\varrho^\delta$.*

Preuve. Si $\chi(\varrho)=0$, il existe $x\in E$ tel que $\varrho x=u(x)$, $x\neq 0$; d'où, puisque $u^\nu=0$, $\varrho^\nu x=0$, c'est-à-dire $\varrho=0$; $\chi(\varrho)$ se réduit donc à son terme de plus haut degré, qui est évidemment ϱ^δ.

Le lemme 10. 1 a pour conséquence évidente le lemme que voici:

Lemme 10. 2. *Soit v un endomorphisme d'un espace euclidien E de dimension δ; si $u=\lambda-v$ est nilpotent, la fonction caractéristique de v est $\chi(\varrho)=(\varrho-\lambda)^\delta$.*

Le lemme 10. 2 permet de résumer et compléter comme suit les lemmes du n⁰ 9:

Théorème 10. 1. ([7], Satz 1', 2, 6, 8, 9, 10, 11, 13) *Soit λ une valeur propre de v autre que* 0.

a) *L'ensemble des vecteurs principaux de v qui correspondent à λ est un sous-espace fermé et euclidien Y de X. La restriction v_Y de v à Y est un endomorphisme de Y; la fonction caractéristique de v_Y est*

$$\chi(\varrho)=(\varrho-\lambda)^\delta \quad \textit{où} \quad \delta=\dim Y; \tag{10. 2}$$

sa seule valeur propre est λ; ses vecteurs principaux et propres sont ceux de v qui correspondent à λ.

b) *Il existe un sous-espace fermé Z de X tel que $v(Z)\subset Z$, $X=Y+Z$, cette somme étant directe. La restriction v_Z de v à Z est un endomorphisme complètement continu de Z; les valeurs propres de v_Z sont les valeurs propres de v autres que λ; à ces valeurs correspondent, pour v_Z et v, les mêmes vecteurs principaux et propres.*

Remarque. On a, vu (10. 1) et (10. 2)

$$u^\delta(Y)=0. \tag{10. 3}$$

Preuve de a). Le lemme 9. 3 prouve que $Y=\overset{-\sigma}{u}(0)$; donc, vu le lemme 9. 2, Y est un sous-espace fermé et euclidien. Soit $u_Y=\lambda-v_Y$; $u_Y^\sigma=0$; d'où, vu le lemme 10. 2, la formule (10. 2), qui prouve que la seule valeur propre de v_Y est λ; les vecteurs principaux et propres correspondant à λ sont les mêmes pour v_Y et v, puisque $\overset{-\nu}{u}(0)\subset Y$ quel que soit $\nu>0$.

Preuve de b). Soit $Z=u^\sigma(X)$; les lemmes 9. 5 et 9. 6 prouvent que Z est fermé et que $X=Y+Z$ (somme directe). Le lemme 9. 5 prouve en outre que λ n'est pas valeur propre de v_Z. Soit $w(x)=\varrho x-v(x)$, où $\varrho\neq\lambda$ posons $x=y+z$, où $y\in Y$, $z\in Z$; l'équation $w^\nu(x)=0$ implique $w^\nu(y)=0$ donc, vu a), $y=0$: les vecteurs principaux et propres, ne correspondant pas à λ, de v et v_Z sont donc les mêmes.

---

[2]) u est nilpotent quand il existe un entier $\nu>0$ tel que $u^\nu=0$.

V. ENDOMORPHISMES TRANSPOSÉS.

11. Notations. X et X^* seront deux espaces vectoriels à voisinages convexes; $\langle x, x^* \rangle$ sera une *application bilinéaire* de leur produit $X \times X^*$ dans A. Nous dirons que la partie Y de X est orthogonale à la partie Z^* de X^* quand $\langle y, z^* \rangle = 0$ quels que soient $y \in Y$, $z^* \in Z^*$. Nous supposerons que 0 *est le seul élément de X (de X^*) orthogonal à X^* (à X)*: X pourra être identifié à un sous-espace du dual ([1], chapitre II, § 4) de X^* et vice-versa; les sous-espaces Y et Y^* de X et X^* seront dits *duals* quand Y s'identifie au dual de Y^* et Y^* au dual de Y.

v et ${}^tv$ seront deux endomorphismes de X et X^* *complètement continus et transposés l'un de l'autre:* $\langle v(x), x^* \rangle = \langle x, {}^tv(x^*) \rangle$ quels que soient $x \in X$ et $x^* \in X^*$.

Soit $\lambda \in A$, $\lambda \neq 0$; le théorème 10. 1 définit des sous-espaces Y et Z de X, si λ est valeur propre; sinon, nous définirons: $Y = 0$, $X = Z$. À ${}^tv$ et λ sont attachés de même deux sous-espaces Y^* et Z^* de X^*.

Lemme 11. 1. *Y et Z^* (Z et Y^*) sont orthogonaux.*

Preuve. $Y = 0$, si λ n'est pas valeur propre de v. Sinon: $u^\delta(Y) = 0$, vu (10. 3); $Z^* = {}^tu(Z^*) = {}^tu^\delta(Z^*)$ vu le théorème 7. 1 ou 10. 1 b; par suite, si $y \in Y$ et $z^* \in Z^*$,

$$\langle y, z^* \rangle = \langle y, {}^tu^\delta(x^*) \rangle = \langle u^\delta(y), x^* \rangle = \langle 0, x^* \rangle = 0.$$

Lemme 11. 2. *Y et Y^* sont duals.*

Preuve. Soit $y \in Y$ orthogonal à Y^*; d'après le lemme 11. 1, y est orthogonal à $Y^* + Z^* = X^*$; donc $y = 0$. Par suite Y est un sous-espace du dual de Y^*; donc, vu [1], ch. II, § 3 proposition 6 et § 4, proposition 6: $\dim Y \leq \dim Y^*$; Y et Y^* sont duals, si l'égalité est réalisée. Or c'est le cas, car, en intervertissant Y et Y^*, on obtient: $\dim Y^* \leq \dim Y$.

Lemme 11. 3. *$u^\nu(Y)$ et ${}^tu^\nu(Y^*)$ sont duals* ($\nu > 0$).

Preuve. [1], chapitre II. § 4, n° 9, th. 4.

Lemme 11. 4. *$\overset{-\nu}{u}(0)$ et ${}^t\overset{-\nu}{u}(0)$ ont même dimension* ($\nu > 0$).

Preuve. On a ([1], ch. II, § 3, proposition 10): $\dim \overset{-\nu}{u}(0) + \dim u^\nu(Y) = \dim Y$; $\dim {}^t\overset{-\nu}{u}(0) + \dim {}^tu^\nu(Y^*) = \dim Y^*$; or Y et $u^\nu(Y)$ ont mêmes dimensions que leurs duals Y^* et ${}^tu^\nu(Y^*)$: ([1], ch. II, § 4, proposition 6).

Lemme 11. 5. *$u(X)$ est l'ensemble des points de X orthogonaux à ${}^t\overset{-1}{u}(0)$.*

Preuve. D'après le théorème 7. 1 ou 10. 1, $u(Z) = Z$ et $u(X) = u(Y) + Z$; $u(Y)$ est l'ensemble des points de Y orthogonaux à ${}^t\overset{-1}{u}(0)$ ([1], ch. II, § 4, th. 3); Z est orthogonal à ${}^t\overset{-1}{u}(0)$ (lemme 11. 1).

Ces cinq lemmes prouvent le théorème suivant:

Théorème 11.1. *Les valeurs propres non nulles des deux endomorphismes v et ${}^t v$ sont les mêmes. Les vecteurs principaux de v et ${}^t v$ correspondant à l'une de ces valeurs propres constituent deux sous-espaces de X et X^*, duals. Les vecteurs propres de v et ${}^t v$ correspondant à l'une de ces valeurs propres constituent deux sous-espaces de X et X^* ayant même dimension. Pour que l'équation d'inconnue x*

$$\lambda x - v(x) = x' \qquad (\lambda \neq 0)$$

ait au moins une solution, il faut et il suffit que x' soit orthogonal aux vecteurs propres de ${}^t v$ qui correspondent à λ.

12. Problème. *Soit ${}^t v$ le transposé, dans le dual topologique de X, d'un endomorphisme complètement continu v de X; pour que ${}^t v$ soit complètement continu, est il nécessaire et suffisant que v le soit?* J. SCHAUDER [8] l'a prouvé quand X est un espace de Banach.

Nota. Le dual topologique de X est l'espace vectoriel à voisinages convexes que constituent les applications linéaires et continues de X dans A. Par définition 0 est le seul point du dual topologique de X orthogonal à X; 0 est le seul point de X orthogonal au dual topologique de X (conséquence aisée de [1], ch. II, § 3, th. 2). Si la réponse à la question posée est affirmative, le théorème 11.1 peut être appliqué, quels que soient X et v, au dual topologique X^* de X et au transposé ${}^t v$ de v dans X^*.

Bibliographie.

[1] N. BOURBAKI, *Éléments de mathématique.* Livre II, ch. II: *Algèbre linéaire* (Paris, 1947).

[2] N. BOURBAKI, *Éléments de mathématique.* Livre III, ch. I: *Structures topologiqus* (Paris, 1940).

[3] J. LERAY et J. SCHAUDER, Topologie et équations fonctionnelles, *Annales de l'École Normale Supérieure,* 51 (1934). p. 45–78.

[4] J. LERAY, Topologie des espaces abstraits de Banach, *Comptes Rendus Acad. Sci. Paris,* **200** (1935), p. 1082–1084.

[5] J. LERAY, Sur les équations et les transformations, *Journal de Math.,* **24** (1945), p. 201–248.

[6] J. LERAY, Théorie des équations dans les espaces vectoriels à voisinages convexes (à paraître; professé au Collège de France en 1948–49; étend [3] et [4] aux espaces vectoriels à voisinages convexes).

[7] F. RIESZ, Über lineare Funktionalgleichungen, *Acta Math.,* **41** (1918), p. 71–98.

[8] J. SCHAUDER, Über lineare, vollstetige Funktionaloperationen, *Studia Math.,* **2** (1930), p. 183–196.

[9] B. L. VAN DER WAERDEN, *Moderne Algebra,* vol. 2 (Berlin, 1940).

[10] A. WEIL, *L'intégration dans les groupes topologiques et ses applications* (Paris, 1940).

(Reçu le 29 août 1949)

The asymptotic behaviour of the coefficients of certain power series.

By G. SZEKERES in Adelaide (Australia).

1. The purpose of this paper is to derive an asymptotic expansion (as $k\to\infty$) for the coefficient $A_k = A_k(K, \alpha)$ in

(1) $$g(z) = e^{K[z-f(\alpha z)]} = \sum_{\mu=0}^{\infty} \frac{1}{\mu!} A_\mu z^\mu, \quad K = \frac{1}{\sigma} k + \tau,$$

where σ, τ are real constants, $\sigma > 0$, and $f(z)$ is analytic at $z=0$.

Under certain assumptions on $f(z)$ one can easily obtain a principal term for A_k by a method introduced by P. DEBYE[1]), called the method of steepest descent (Sattelpunktmethode). Using the theorem of residues we have

$$\frac{1}{k!} A_k = \frac{1}{2\pi i} \int \frac{1}{z} e^{\left(\frac{1}{\sigma}k+\tau\right)[z-f(\alpha z)] - k \log z} dz$$

taken over a small circle surrounding the origin. By CAUCHY's theorem we can displace the contour so that it shall pass through the "saddle point" z_0 at which $\frac{d}{dz}\left(\frac{1}{\sigma}(z-f(\alpha z)) - \log z\right) = 0$ or, $z_0 = \sigma + \alpha z_0 f'(\alpha z_0)$. It can be shown that if the contour passes through z_0 in a suitable direction and $f(z)$ behaves appropriately on the rest of the path of integration, then the chief contribution to the integral is furnished by the neighbourhood of z_0 and has the form

(2) $$(2\pi c k)^{-\frac{1}{2}} e^{\left(\frac{1}{\sigma}k+\tau\right)[z_0 - f(\alpha z_0)] - k \log z_0}$$

for a certain constant c which depends on the particular conditions of the problem. The difficulty of the method lies in the proper choice of the path of integration which has to be determined individually for each function $f(z)$. If we cannot make sure that the contribution due to the complementary part of the path is negligible then the method has a rather heuristic value.

In the present paper I shall obtain a full asymptotic expansion for A_k by an entirely different method which avoids complex integration altogether.

---

[1]) See e. g. G. SZEGŐ, *Orthogonal polynomials* (New York, 1939), p. 215 ff. I am indebted to P. TURÁN for having called my attention to this method.

The method has a more or less formal elementary character and does not require a knowledge of the functiontheoretical properties of $f(z)$ except that it has a power series at $z=0$. In fact, I shall prove the following

Theorem 1. *Let* $f(z)=\sum_{\nu=1}^{\infty}\frac{1}{\nu}c_\nu z^\nu$ *be an arbitrary power series with positive radius of convergence and constant term zero*[2]*;* A_k *shall be defined as above. Let* $u=u(t)=\sum_{\nu=0}^{\infty}d_\nu t^{\nu+1}$, $d_0=\sigma$, *denote the inverse function of*

$$t=\frac{u}{\sigma+uf'(u)},\quad u=t\left(\sigma+\sum_{\nu=1}^{\infty}c_\nu u^\nu\right) \tag{3}$$

and let $v=v(t)=\sum_{\nu=0}^{\infty}D_\nu t^{\nu+1}$, $D_0=\sigma$ *be defined by*

$$v=t\left(\sigma+\sum_{\nu=1}^{\infty}|c_\nu|v^\nu\right). \tag{4}$$

Then

$$A_k=K^k\left(\frac{u(\alpha)}{\sigma\alpha}\right)^{\sigma\tau-1}\left(\frac{1}{\sigma}u'(\alpha)\right)^{\frac{1}{2}}\exp\left\{-K\int_0^\alpha\frac{u(t)-\sigma t}{t^2}\,dt\right\}\left(1+\sum_{\mu=1}^{m}k^{-\mu}\psi_\mu(\alpha)+O(k^{-m-1})\right) \tag{5}$$

for every $m\geqq 0$ *and certain functions* $\psi_\mu(\alpha)$, *analytic in* $|\alpha|<\varrho$, *where* ϱ *denotes the radius of convergence of* $v(t)$. *The expansion (5) is uniformly valid for* $|\alpha|\leqq\varrho_0<\varrho$.

The constant in $O(k^{-m-1})$ depends, of course, on ϱ_0 and on m and a similar remark pertains to all O-notations in the paper. $\left(\frac{u}{\sigma\alpha}\right)^{\sigma\tau-1}$ and $\left(\frac{1}{\sigma}u'\right)^{\frac{1}{2}}$ denote the principal branches of the functions, which have the value 1 at $\alpha=0$. No ambiguity is involved in $\int_0^\alpha\frac{u-\sigma t}{t^2}dt$ for complex values of α since the integrand is obviously regular for $|t|\leqq|\alpha|$.

The range of validity of (5) is not necessarily confined to the circle with radius ϱ. In fact, it might be possible to "continue" the expansion into new regions of the complex plane by varying the coefficient c_1. It is quite possible that (5) always holds in the interior of the circle of convergence of $u(t)$.

Using Stirling's formula, we obtain from (5) for the coefficients of the power series (1)

$$\begin{aligned}\frac{1}{k!}A_k=(2\pi)^{-\frac{1}{2}}e^{\sigma\tau}\left(\frac{u(\alpha)}{\sigma\alpha}\right)^{\sigma\tau-1}\left(\frac{1}{\sigma}u'(\alpha)\right)^{\frac{1}{2}}k^{-\frac{1}{2}}\left(\frac{e}{\sigma}\right)^k\times\\ \times\exp\left\{-K\int_0^\alpha\frac{u-\sigma t}{t^2}dt\right\}\left(1+\sum_{\mu=1}^{m}k^{-\mu}\lambda_\mu(\alpha)+O(k^{-m-1})\right)\end{aligned} \tag{6}$$

---

[2]) Obviously this last assumption can be made without loss of generality.

where the $\lambda_\mu(\alpha)$ are analytic for $|\alpha| < \varrho$. The principal term in (6) is identical with (2) if we put $z_0 = \frac{u(\alpha)}{\alpha}$, $c = 1 - \frac{1}{\sigma}\alpha^2 z_0^2 f''(\alpha z_0)$ since

$$-\int_0^\alpha \frac{u(t)-\sigma t}{t^2}\,dt = \frac{u(\alpha)}{\alpha} - f(u(\alpha)) - \sigma \log\frac{u(\alpha)}{\sigma\alpha} - \sigma.$$

To have a simple application put $\sigma = \frac{1}{2}$, $\tau = 1$, $f(z) = \frac{1}{4}z^2$, $u(t) = \frac{1-\sqrt{1-t^2}}{t}$, $\alpha = (\cosh\varphi)^{-1}$. Obviously $A_k = \left(\frac{(2k+1)^{1/2}}{2\cosh\varphi}\right)^k H_k(x)$, $x = (2k+1)^{\frac{1}{2}}\cosh\varphi$, where $H_k(x)$ is the k-th Hermite polynomial. We have

$$-\int_0^\alpha \frac{u-\frac{1}{2}t}{t^2}\,dt = \frac{1}{2}\left(\frac{1}{\alpha^2}(1-\sqrt{1-\alpha^2}) + \log\frac{1+\sqrt{1-\alpha^2}}{2} - \frac{1}{2}\right) =$$

$$= \frac{1}{2}\left(\varphi + \frac{1}{2}e^{-2\varphi} - \log(2\cosh\varphi)\right),$$

$$u'(\alpha) = (1-\sqrt{1-\alpha^2})\,\alpha^{-2}(1-\alpha^2)^{-\frac{1}{2}} = \frac{u(\alpha)}{\alpha}\coth\varphi,$$

hence from (5)

$$A_k \sim (2k+1)^k (2\sinh\varphi)^{-\frac{1}{2}} (2\cosh\varphi)^{-k} \exp\left\{\left(k+\frac{1}{2}\right)\left(\varphi + \frac{1}{2}e^{-2\varphi}\right)\right\} \times$$
$$\times (1 + k^{-1}\psi_1(\varphi) + k^{-2}\psi_2(\varphi) + \ldots),$$

$$e^{-\frac{1}{2}x^2} H_k(x) = 2^{\frac{k-1}{2}} \left(\frac{k}{e}\right)^{\frac{1}{2}k} (\sinh\varphi)^{-\frac{1}{2}} \exp\left\{\left(k+\frac{1}{2}\right)\left(\varphi - \frac{1}{2}\sinh 2\varphi\right)\right\} \times$$
$$\times \left(1 + \sum_{\mu=1}^{m} k^{-\mu}\psi_\mu(\varphi) + O(k^{-m-1})\right)$$

for $x = (2k+1)^{\frac{1}{2}}\cosh\varphi$, $\varphi \geq \varepsilon > 0$. This is a well known result due to PLANCHEREL and ROTACH[3]), who obtained it for real values of φ. The above proof is valid for any complex φ with $|\cosh\varphi| \geq \varepsilon$. The corresponding result for the "oscillating region" $x = (2n+1)^{\frac{1}{2}}\cos\varphi$, $\varepsilon \leq \varphi \leq \pi - \varepsilon$ cannot be obtained by our method.

There is another important application of Theorem 1 to which I intend to come back in another paper, namely, the problem of the asymptotic evaluation of certain partition functions such as the number of partitions of an integer n into exactly k parts. They depend on asymptotic developments of the same type as discussed in the present paper.

---

[3]) M. PLANCHEREL—W. ROTACH, Sur les valeurs asymptotiques des polynomes d'Hermite, *Commentarii Math. Helvetici,* **1** (1929), pp. 227—254. These authors use the method of steepest descent in their work. Regarding above form of the formula (with $m = 0$) see G. SZEGŐ, l. c., p. 195.

2. Proof of the theorem. We have from (1),

$$g'(z)=K\left(1-\sum_{\nu=1}^{\infty} c_\nu \alpha^\nu z^{\nu-1}\right)g(z)$$

hence

$$A_0=1,\ A_1=K(1-c_1\alpha),$$

(7)
$$A_p=K(1-c_1\alpha)A_{p-1}-K\sum_{\nu=2}^{p}(p-1)\dots(p-\nu+1)c_\nu\alpha^\nu A_{p-\nu} \quad \text{for } p>1.$$

We assume that $A_p \neq 0$ for $p=1,\dots,k$ [4]) and write

(8)
$$\varphi_s=s\frac{A_{s-1}}{A_s} \quad \text{for } s>0, \quad \varphi_s=0 \quad \text{for} \quad s\leq 0.$$

Then

(9)
$$A_p^{-1}=\frac{1}{p!}\varphi_1\dots\varphi_p$$

and by (7) and (8)

(10)
$$\varphi_p=\frac{p}{K}+\sum_{\nu=1}^{\infty}c_\nu\alpha^\nu\varphi_p\dots\varphi_{p-\nu+1} \quad \text{for } p>0.$$

From (9), $-\log A_k=-\log k!+\sum_{p=1}^{k}\log\varphi_p$, hence

(11)
$$-\frac{d}{d\alpha}(\log A_k(\alpha))=\sum_{p=1}^{k}\frac{1}{\varphi_p(\alpha)}\frac{d}{d\alpha}\varphi_p(\alpha).$$

The right hand side of (11) can be expressed in a very simple manner by $\varphi_k(\alpha)$ in consequence of the following differential — difference equation:

(12)
$$\frac{t}{\varphi_p(t)}\frac{d}{dt}\varphi_p(t)=K(\varphi_p-\varphi_{p-1})-1, \quad p\geq 1.$$

The formula is obviously true for $p=1$, since $\varphi_0(t)=0$, $\varphi_1(t)=\frac{1}{K}(1-c_1t)^{-1}$, hence we may assume its validity for $1\leq i<p$. Summing (12) for $i=1,\dots,r$ we obtain

(13)
$$\sum_{i=1}^{r}t\varphi_i^{-1}\varphi_i'=K\varphi_r-r \quad \text{for} \quad r<p.$$

Let us write $\pi_\nu=\prod_{i=1}^{\nu-1}\varphi_{p-i}$. Then from (10)

$$t\varphi_p^{-1}\varphi_p'=\sum_{\nu=1}^{p}\nu c_\nu t^\nu\pi_\nu+\sum_{\nu=1}^{p}c_\nu t^\nu\pi_\nu\sum_{i=0}^{\nu-1}t\varphi_{p-i}^{-1}\varphi_{p-i}'$$
$$=\sum_{\nu=1}^{p}\nu c_\nu t^\nu\pi_\nu+\sum_{\nu=1}^{p}c_\nu t^\nu\pi_\nu t\varphi_p^{-1}\varphi_p'+\sum_{\nu=1}^{p}c_\nu t^\nu\pi_\nu(K(\varphi_{p-1}-\varphi_{p-\nu})-(\nu-1))$$

by the induction hypothesis, hence

---

[4]) This assumption in convenient but not really essential for the proof. We can always enforce the condition by slightly changing the value of the parameter α.

$$\left(1-\sum_{\nu=1}^{p} c_\nu t^\nu \pi_\nu\right) t\varphi_p^{-1}\varphi_p' = \sum_{\nu=1}^{p} c_\nu t^\nu \pi_\nu (K(\varphi_{p-1}-\varphi_{p-\nu})+1) =$$

$$= \sum_{\nu=1}^{p} c_\nu t^\nu \pi_\nu (K(\varphi_p-\varphi_{p-\nu}) - K(\varphi_p-\varphi_{p-1})+1) =$$

$$= K\sum_{\nu=1}^{p} c_\nu t^\nu (\varphi_p\varphi_{p-1}\dots\varphi_{p-\nu+1}-\varphi_{p-1}\dots\varphi_{p-\nu}) - \sum_{\nu=1}^{p} c_\nu t^\nu \pi_\nu (K(\varphi_p-\varphi_{p-1})-1) =$$

$$= K\left[\left(\varphi_p-\frac{p}{K}\right)-\left(\varphi_{p-1}-\frac{p-1}{K}\right)\right] - \sum_{\nu=1}^{p} c_\nu t^\nu \pi_\nu (K(\varphi_p-\varphi_{p-1})-1) =$$

$$= \left(1-\sum_{\nu=1}^{p} c_\nu t^\nu \pi_\nu\right)(K(\varphi_p-\varphi_{p-1})-1)$$

whence our assertion follows $\left(\text{since } (1-\sum_{\nu=1}^{p} c_\nu t^\nu \pi_\nu)\,\varphi_p = \frac{p}{K} \neq 0\right)$. Hence (13) is valid for $r=p$, $\sum_{i=1}^{p}\varphi_i^{-1}\varphi_i' = \frac{K}{t}\left(\varphi_p-\frac{p}{K}\right)$ for $p=1,2,\dots$ and $-\frac{d}{d\alpha}\log A_k(\alpha) = \frac{K}{\alpha}\left(\varphi_k(\alpha)-\frac{k}{K}\right)$ by (11). Integrating and noting that, by (9), $\varphi_p(0)=\frac{p}{K}$ and $A_k(0)=K^k$, we obtain

(14) $$A_k = K^k \exp\left\{-K\int_0^\alpha \left(\varphi_k(t)-\frac{k}{K}\right)t^{-1}dt\right\}.$$

This relation reduces the problem to the asymptotic evaluation of $\varphi_k(t)$. Write $\varphi_p(t)=\sum_{\nu=0}^{\infty} d_p(\nu)t^\nu$ where $d_p(\nu)=0$ for $p\leq 0$, $d_p(0)=\frac{p}{K}$ for $p>0$, and

(15) $$d_p(1)t+d_p(2)t^2+\dots = \sum_{\nu=1}^{\infty} c_\nu t^\nu \prod_{r=1}^{\nu}\left(\sum_{i=0}^{\infty} d_{p-r+1}(i)t^i\right) \text{ for } p>0.$$

The idea naturally suggests itself to compare $\varphi_p(t)$ with the function $\varphi=\varphi\left(\frac{p}{K},t\right)$ defined by the equation $\varphi=\frac{p}{K}+\sum_{\nu=1}^{\infty} c_\nu t^\nu \varphi^\nu$. Generally let

(16) $$\varphi(\xi,t)=\sum_{\nu=0}^{\infty} d(\xi,\nu)\,t^\nu, \quad d(\xi,0)=\xi$$

be defined for $\xi>0$ by

(17) $$\varphi(\xi,t)=\xi+\sum_{\nu=1}^{\infty} c_\nu t^\nu \varphi^\nu(\xi,t)=\xi+t\varphi(\xi,t)f'(t\varphi(\xi,t)).$$

From (16) we have

(18) $$d(\xi,1)t+d(\xi,2)t^2+\dots=\sum_{\nu=1}^{\infty} c_\nu t^\nu[\xi+d(\xi,1)t+d(\xi,2)t^2+\dots]^\nu$$

whence $d(\xi,1)=c_1\xi$, $d(\xi,2)=c_2\xi^2+c_1^2\xi,\dots,$

(19) $$d(\xi,\nu)=\sum_{r_1+\dots+r_i=\nu} b(r_1,\dots,r_i)\,c_{r_1}\dots c_{r_i}\,\xi^{\nu-i+1} \quad \text{for } \nu>0$$

where the $b(r_1, \ldots, r_i)$ are positive absolute constants and the summation runs over the (unrestricted) partitions of ν.

Comparing (17) and (3) we see that

(20) $u(t) = t\varphi(\sigma, t), \quad d_\nu = d(\sigma, \nu) = \sum_{r_1+\cdots+r_i=\nu} b(r_1, \ldots, r_i)\, c_{r_1} \ldots c_{r_i} \sigma^{\nu-i+1}$

and $|d_\nu| \leqq \sum b(r_1, \ldots, r_i)\, |c_{r_1} \ldots c_{r_i}|\, \sigma^{\nu-i+1} = D_\nu$. This shows that $|d(\xi, \nu)| \leqq D_\nu$ if $\xi \leqq \sigma$ and $|d(\xi, \nu)| \leqq \left(\frac{\xi}{\sigma}\right)^\nu D_\nu$ if $\xi > \sigma$. Hence, if $\varrho(\xi)$ denotes the radius of convergence of (16), then $\varrho(\xi) \geqq \varrho$ if $\xi \leqq \sigma$ and $\varrho(\xi) \geqq \varrho \frac{\sigma}{\xi}$ if $\xi > \sigma$. In particular

$$\varrho\left(\frac{k}{K}\right) = \varrho\left(\frac{k}{\frac{1}{\sigma} k + \tau}\right) = \varrho\left(\left(1 - \frac{\tau}{\frac{1}{\sigma} k + \tau}\right)\sigma\right) \geqq \left(1 - \frac{\sigma|\tau|}{k}\right)\varrho,$$

and $\varphi\left(\frac{k}{K}, t\right)$ is certainly convergent for sufficiently large k if $|t| \leqq \varrho_0 < \varrho$. (We only have to take $k > \frac{\varrho}{\varrho - \varrho_0} \sigma|\tau|$.) It follows that $\varphi\left(\frac{p}{K}, \alpha\right)$ is convergent for every $p \leqq k$. Comparing the coefficients $d_p(\nu)$ and $d\left(\frac{p}{K}, \nu\right)$ for $p \leqq k$, $\nu = 0, 1, \ldots$ we find from (15) and (18) $d_p(0) = d\left(\frac{p}{K}, 0\right) = \frac{p}{K}$,

$$d_p(1) = d\left(\frac{p}{K}, 1\right) = c_1 \frac{p}{K}, \quad d_p(2) = \frac{p}{K} c_1^2 + \frac{p(p-1)}{K^2} c_2 = d\left(\frac{p}{K}, 2\right) - \frac{p}{K^2} c_2,$$

$d_p(\nu) = d\left(\frac{p}{K}, \nu\right) + O\left(\frac{h(\nu)}{K}\right)$ for every fixed ν. This shows that the error commited by taking $\varphi\left(\frac{k}{K}, t\right)$ instead of $\varphi_k(t)$ is small; in fact, $\varphi_r(\alpha) = \varphi\left(\frac{p}{K}, \alpha\right) + O\left(\frac{1}{K}\right)$ provided that $h(\nu)$ is not increasing too rapidly with ν. In order to have an estimate for $h(\nu)$ we introduce the following notation:

For $r > 0$, $i \geqq 0$ write

$$(1 - ix)(1 - (i+1)x) \ldots (1 - (i+r-1)x) \equiv \sum_{\nu=0}^{r} (-1)^\nu S(i, r; \nu) x^\nu,$$

where $S(i, r; 0) = 1$, and denote by $P_p(i, r)$, $p > i$ an arbitrary expression

(21) $$P_p(i, r) = \sum_{\nu=0}^{\infty} (-1)^\nu \alpha_\nu p^{-\nu}$$

where $\alpha_0 = 1$ and the coefficients α_ν satisfy for $\nu > 0$

(22) $$0 \leqq \alpha_\nu \leqq S(i, r; \nu)$$

(hence $\alpha_\nu = 0$ for $\nu > r$), and

(23) $$\sum_{\nu=0}^{\infty} (-1)^\nu \alpha_{s+\nu} p^{-s-\nu} \geqq 0 \quad \text{for } s = 0, 1, \ldots, r.$$

The second condition clearly implies

(24) $$\sum_{\nu=0}^{\infty}(-1)^{\nu}\alpha_{s+\nu}p^{-s-\nu}\leqq\alpha_s p^{-s} \quad \text{for} \quad s=0,\ldots,r.$$

If $a_1>0$, $a_2>0$ then

(25) $$a_1P_p(i,r)+a_2P_p(i,r)=(a_1+a_2)P_p(i,r).$$

In this formula (like in others which follow) each $P_p(i,r)$ may denote a different expression of the form (21). The formula is to be read from the left to right: if $a_1>0$, $a_2>0$ and both expressions $P_p(i,r)$ on the left satisfy (22) and (23) then the $P_p(i,r)$ on the right (which is uniquely determined by the left hand side) also satisfies these conditions. The proof is obvious

The following formula is less trivial:

(26) $$P_p(i,r)P_p(j,s)=P_p(i,r+s) \quad \text{if} \quad j\leqq i+r.$$

Write $P_p(i,r)=\sum(-1)^{\nu}\alpha_{\nu}p^{-\nu}$, $P_p(j,s)=\sum(-1)^{\nu}\beta_{\nu}p^{-\nu}$ and $P_p(i,r)\,P_p(j,s)=\sum(-1)^{\nu}\gamma_{\nu}p^{-\nu}$ where $\gamma_{\nu}=\alpha_0\beta_{\nu}+\ldots+\alpha_{\nu}\beta_0$.

Comparing $(1-ix)\ldots(1-(i+r-1)x)(1-jx)\ldots(1-(j+s-1)x)$ with $(1-ix)\ldots(1-(i+r+s-1)x)$ we see that $\sum_{\mu=0}^{\nu}S(i,r;\mu)S(j,s;\nu-\mu)\leqq$ $\leqq S(i,r+s;\nu)$ if $j\leqq i+r$ hence $\gamma_{\nu}=\alpha_0\beta_{\nu}+\ldots+\alpha_{\nu}\beta_0\leqq S(i,r+s;\nu)$ which proves condition (22). Condition (23) is a consequence of the following

Lemma. *Let $\{a_0,a_1,\ldots\}$, $\{b_0,b_1,\ldots\}$ be non-negative sequences with finite sums. Let us form the alternating series $\sum(-1)^{\nu}a_{\nu}$, $\sum(-1)^{\nu}b_{\nu}$ and suppose that the remainders of the two series themselves form an alternating sequence, i. e. $a_{\nu}-a_{\nu+1}+a_{\nu+2}-\ldots\geqq 0$, $b_{\nu}-b_{\nu+1}+b_{\nu+2}-\ldots\geqq 0$ for $\nu\geqq 0$. Then the remainders of the Cauchy product $c_{\nu}=a_0b_{\nu}+\ldots+a_{\nu}b_0$ also form an alternating sequence: $c_{\nu}-c_{\nu+1}+c_{\nu+2}-\ldots\geqq 0$.*

The lemma is notably true if both $\{a_{\nu}\}$ and $\{b_{\nu}\}$ are finite.

Proof. $$\sum_{r=0}^{\infty}(-1)^r c_{\nu+r}=\left(\sum_{r=0}^{\infty}(-1)^r a_r\right)\left(\sum_{r=0}^{\infty}(-1)^r b_{\nu+r}\right)+$$
$$+b_{\nu-1}\sum_{r=0}^{\infty}(-1)^r a_{1+r}+b_{\nu-2}\sum_{r=0}^{\infty}(-1)^r a_{2+r}+\ldots+b_0\sum_{r=0}^{\infty}(-1)^r a_{\nu+r}\geqq 0$$

since each term is $\geqq 0$.

Applying the lemma to $a_{\nu}=\alpha_{\nu}p^{-\nu}$, $b_{\nu}=\beta_{\nu}p^{-\nu}$, $c_{\nu}=\gamma_{\nu}p^{-\nu}$, we immediately obtain condition (23) for the right hand side of (26). A repeated application of the lemma also shows that

$$\left(1-\frac{i}{p}\right)\ldots\left(1-\frac{i+r-1}{p}\right)\equiv\sum_{\nu=0}^{r}(-1)^{\nu}p^{-\nu}S(i,r;\nu)=P_p(i,r).$$

We shall now show that for $p > i \geqq 0$, $\nu > 0$

$$d_{p-i}(\nu) = \sum b(r_1, \ldots, r_i) c_{r_1} \ldots c_{r_i} \left(\frac{p}{K}\right)^{\nu-i+1} P_p(i, \nu) \tag{27}$$

summed for the partitions $r_1 + \ldots + r_i = \nu$. For the expression on the right we use the shorthand notation $d\left(\frac{p}{K}, \nu\right) \times P_p(i, \nu)$, the symbol $\times$ reminding us of the fact that $d\left(\frac{p}{K}, \nu\right)$ is a sum of terms (19) each term being multiplied by a $P_p(i, \nu)$.

Writing $p-i$ instead of p in (15), we obtain by comparing the coefficients of t, $d_{p-i}(1) = c_1 \frac{p-i}{K} = c_1 \frac{p}{K}\left(1 - \frac{i}{p}\right) = d\left(\frac{p}{K}, 1\right) P_p(i, 1)$ for $i < p$ which proves (27) for $\nu = 1$. We note that $d_{p-i}(0) = \frac{p-i}{K} = d\left(\frac{p}{K}, 0\right) P_p(i, 1)$, hence assuming the induction hypothesis,

$$d_{p-i}(r) = d\left(\frac{p}{K}, r\right) \times P_p(i, r+1) \quad \text{for} \quad i < p, \quad 0 \leqq r < \nu.\,^{5)} \tag{28}$$

Equating the coefficients of t^ν in (15) and writing again $p-i$ in the place of p, we obtain

$$d_{p-i}(\nu) = \sum_{\mu=1}^{\nu} c_\mu \sum_{s_1+\ldots+s_\mu=\nu-\mu} d_{p-i}(s_1)\, d_{p-i-1}(s_2) \ldots d_{p-i-\mu+1}(s_\mu), \tag{29}$$

the summation extending over every composition of $\nu-\mu$ with $s_1, \ldots, s_\mu \geqq 0$. Since $s_1 + \ldots + s_\mu < \nu$, we have by (28) for every non-zero term of the sum

$$\begin{aligned} d_{p-i}(s_1) \ldots d_{p-i-\mu+1}(s_\mu) &= \\ = d\left(\frac{p}{K}, s_1\right) \ldots d\left(\frac{p}{K}, s_\mu\right) \times P_p(i, s_1+1) P_p(i+1, s_2+1) \ldots P_p(i+\mu-1, s_\mu+1) &= \\ = d\left(\frac{p}{K}, s_1\right) \ldots d\left(\frac{p}{K}, s_\mu\right) \times P_p(i, \nu) & \end{aligned} \tag{30}$$

since

$$\begin{aligned} P_p(i, s_1+1) P_p(i+1, s_2+1) \ldots P_p(i+\mu-1, s_\mu+1) &= \\ = P_p(i, s_1+s_2+2) P_p(i+2, s_3+1) \ldots P_p(i+\mu-1, s_\mu+1) &= \\ = \ldots = P_p(i, s_1+s_2+\ldots+s_\mu+\mu) = P_p(i, \nu) & \end{aligned}$$

by a repeated application of (26). The same is true if $p-i-\mu+1 \leqq 0$, that is, the left hand side of (30) is zero, since then $p < i+\mu \leqq i+\nu$ and obviously $0 = P_p(i, \nu)$ if $i < p < i+\nu$. Now from (18) we have for $\xi = p/K$

$$d\left(\frac{p}{K}, \nu\right) = \sum_{\mu=1}^{\nu} c_\mu \sum d\left(\frac{p}{K}, s_1\right) \ldots d\left(\frac{p}{K}, s_\mu\right) \tag{31}$$

where the summation is to be taken for the same compositions as in (29).

---

5) Here we have used the obvious equation $P_p(i, r) = P_p(i, r+1)$ which should be read from the left to the right.

This shows that to each term $T=c_\mu d\left(\frac{p}{K}, s_1\right)\dots d\left(\frac{p}{K}, s_\mu\right)$ in (31) there is a corresponding term T^* in (29) with $T^*=T\times P_p(i,\nu)$ by (30). The precise meaning of this relation is that if we substitute (19) $\left(\text{with } \xi=\frac{p}{K}\right)$ for each $d\left(\frac{p}{K}, s_i\right)$ in T and carry out term-by-term multiplication then each of these terms appears in T^* multiplied by a $P_p(i,\nu)$. Collecting terms belonging to the same $c_{r_1}\dots c_{r_i}$ we have by (25), since the multiplying constants $b(r_1,\dots,r_i)$ are positive, $d_{p-i}(\nu)=d\left(\frac{p}{K},\nu\right)\times P_p(i,\nu)$, i. e. (27). In particular, $d_p(\nu)=d\left(\frac{p}{K},\nu\right)\times P_p(0,\nu)$ and $d_k(\nu)=d\left(\frac{k}{K},\nu\right)\times P_k(0,\nu)$. Therefore

$$
(32)\quad \begin{aligned}
\varphi_k(t) &= \sum_{\nu=0}^{\infty} d_k(\nu)t^\nu=\sum_{\nu=0}^{\infty} d\left(\frac{k}{K},\nu\right)\times P_k(0,\nu)\,t^\nu=\\
&=\sum_{\nu=0}^{\infty}\sum_{r_1+\dots+r_i=\nu} b(r_1,\dots,r_i)\,c_{r_1}\dots c_{r_i}\left(\frac{k}{K}\right)^{\nu-i+1}(1-k^{-1}\beta(r_1,\dots,r_i;1)+\\
&\qquad+\dots+(-1)^\nu k^{-\nu}\beta(r_1,\dots,r_i;\nu))\,t^\nu
\end{aligned}
$$

where

$$(33)\qquad 0\leqq\beta(r_1,\dots,r_i;j)\leqq S(0,\nu;j)<\nu^{2j},$$

$$(34)\qquad 0\leqq 1-k^{-1}\beta(r_1,\dots,r_i;1)+\dots+(-1)^\nu k^{-\nu}\beta(r_1,\dots,r_i;\nu)\leqq 1.$$

This shows first, that the radius of $\varphi_k(t)$ is not less than $\left(1-\frac{\sigma|\tau|}{k}\right)\varrho$, hence $\varphi_k(\alpha)$ is convergent if k is sufficiently large. Secondly,

$$
(35)\quad \begin{aligned}
\varphi_k(t)&=\varphi\left(\frac{k}{K},t\right)+\\
&+\sum_{\mu=1}^{m}(-1)^\mu k^{-\mu}\sum_{\nu=\mu}^{\infty}\left(\sum_{r_1+\dots+r_i=\nu} b(r_1,\dots,r_i)\,c_{r_1}\dots c_{r_i}\left(\frac{k}{K}\right)^{\nu-i+1}\beta(r_1,\dots,r_i;\mu)\right)t^\nu+\\
&+\sum_{\nu=m+1}^{\infty}\left\{\sum_{r_1+\dots+r_i=\nu} b(r_1,\dots,r_i)\,c_{r_1}\dots c_{r_i}\left(\frac{k}{K}\right)^{\nu-i+1}\left(\sum_{\mu=m+1}^{\nu}(-1)^\mu k^{-\mu}\beta(r_1,\dots,r_i;\mu)\right)\right\}t^\nu
\end{aligned}
$$

for $|t|\leqq\varrho_0$. For, each of the series

$$\sum_{\nu=\mu}^{\infty}\left(\sum b(r_1,\dots,r_i)\,c_{r_1}\dots c_{r_i}\left(\frac{k}{K}\right)^{\nu-i+1}\beta(r_1,\dots,r_i;\mu)\right)t^\nu,\quad \mu=1,\dots,m,$$

is absolutely convergent for $|t|\leqq\varrho_0$ by (33), and also the remainder since

$$\left|\sum_{\mu=m+1}^{\nu}(-1)^\mu k^{-\mu}\beta(r_1,\dots,r_i;\mu)\right|\leqq k^{-m-1}\beta(r_1,\dots,r_i;m+1)<k^{-m-1}\nu^{2m+2}$$

by (24). Hence the remainder term in (35) can be written in the form $k^{-m-1}t\chi(t)$, where $\chi(t)$ is absolutely convergent for $|t|\leqq\varrho_0$.

Finally we replace $\frac{k}{K}$ by σ, $\varphi\left(\frac{k}{K},t\right)$ by $\varphi(\sigma,t)$ in (35) by putting

$$\frac{k}{K}=\frac{k}{\frac{1}{\sigma}k+\tau}=\sigma\left(1-\frac{\sigma\tau}{k}+\ldots+(-1)^m\frac{\sigma^m\tau^m}{k^m}+(-1)^{m+1}\frac{\sigma^{m+1}\tau^{m+1}}{k^m(k+\sigma\tau)}\right)$$ and collecting terms belonging to equal powers of k. As a result we obtain the asymptotic development

$$t^{-1}\left(\varphi_k(t)-\frac{k}{K}\right)=t^{-1}(\varphi(\sigma,t)-\sigma)+\sum_{\mu=1}^{m}k^{-\mu}\psi^*_\mu(t)+O(k^{-m-1}) \tag{36}$$

valid for $|t|\leqq\varrho_0$, hence for $t=\alpha$.

(14), (36) and (20) give

$$A_k=K^k\exp\left\{-K\int_0^\alpha\frac{u-\sigma t}{t^2}\,dt-\frac{1}{\sigma}\int_0^\alpha\psi^*_1(t)\,dt-\right.$$
$$\left.-\sum_{\mu=1}^{m-1}k^{-\mu}\int_0^\alpha\left(\tau\psi^*_\mu+\frac{1}{\sigma}\psi^*_{\mu+1}\right)dt+O(k^{-m})\right\}. \tag{37}$$

It remains to determine the explicit expression for $\psi^*_1(t)$. We have from (36), if we put $u(t)=t\varphi(\sigma,t)$ and $\psi(t)=-t\psi^*_1(t)+\sigma^2\tau$,

$$\varphi_k(t)\sim\frac{k}{K}-\sigma+\frac{u(t)}{t}+\frac{t}{k}\psi^*_1(t)\sim\frac{u(t)}{t}-\frac{1}{k}\psi(t)$$

where $\sim$ indicates that the difference of the left and right hand sides is $O(k^{-2})$. Using equation (12) we obtain

$$\varphi_{k-1}=\varphi_k\quad-\frac{1}{K}(1+t\varphi_k^{-1}\varphi_k')\sim\frac{u}{t}-\frac{1}{k}\psi-\frac{\sigma}{k}t\frac{u'}{u},$$
$$\varphi_{k-2}=\varphi_{k-1}-\frac{1}{K}(1+t\varphi_{k-1}^{-1}\varphi_{k-1}')\sim\frac{u}{t}-\frac{1}{k}\psi-\frac{2\sigma}{k}t\frac{u'}{u},$$

generally $\varphi_{k-i}\sim\frac{u}{t}-\frac{1}{k}\psi-\frac{i\sigma}{k}t\frac{u'}{u}$. We can put these expressions, purely formally, into (10) since we know that the asymptotic expansion (36) is valid

$$\frac{u}{t}-\frac{1}{k}\psi\sim\frac{k}{K}+\sum c_\nu t^\nu\prod_{i=0}^{\nu-1}\left(\frac{u}{t}-\frac{1}{k}\psi-\frac{i\sigma}{k}t\frac{u'}{u}\right)$$
$$\sim\sigma-\frac{\sigma^2\tau}{k}+\sum_{\nu=1}^{\infty}c_\nu u^\nu-\frac{1}{k}\sum_{\nu=1}^{\infty}\nu c_\nu t u^{\nu-1}\psi-\frac{\sigma}{k}\sum_{\nu=1}^{\infty}\binom{\nu}{2}c_\nu t^2u^{\nu-2}u'$$

whence

$$\left(1-\sum_{\nu=1}^{\infty}\nu c_\nu t u^{\nu-1}\right)\psi=\sigma^2\tau+\sigma\sum_{\nu=1}^{\infty}\binom{\nu}{2}c_\nu t^2u^{\nu-2}u'.$$

From (3), $u'=\frac{u}{t}+t\sum\nu c_\nu u^{\nu-1}u'$, $1-\sum\nu c_\nu t u^{\nu-1}=\frac{u}{u't}$,

$\sum\binom{\nu}{2}c_\nu t^2u^{\nu-2}u'=\frac{u}{u't}-1+\frac{1}{2}\frac{uu''}{(u')^2}$, hence $\frac{1}{\sigma}\psi=(\sigma\tau-1)t\frac{u'}{u}+1+\frac{1}{2}t\frac{u''}{u'}$,

$$-\frac{1}{\sigma}\psi^*_1=\frac{1}{t}\left(\frac{1}{\sigma}\psi-\sigma\tau\right)=(\sigma\tau-1)\left(\frac{u'}{u}-\frac{1}{t}\right)+\frac{1}{2}\frac{u''}{u'},$$
$$-\frac{1}{\sigma}\int_0^\alpha\psi^*_1(t)\,dt=(\sigma\tau-1)\log\frac{u(\alpha)}{\sigma\alpha}+\frac{1}{2}\log\frac{u'(\alpha)}{\sigma}.$$

Theorem 1 follows from this and (37).

It would be possible to obtain, in a similar manner, the next term $\psi_2^*(t)$, but the general term $\psi_\mu^*(t)$ could hardly be obtained in that way explicitely. The method of steepest descent gives heuristically, in the case $\sigma=1$, $\tau=1$, the following expressions for $\psi_\mu(\alpha)$ in

$$A_{K-1} \sim K^{K-1}(u'(\alpha))^{\frac{1}{2}} \exp\left\{-K\int_0^\alpha \frac{u-t}{t^2}\,dt\right\}\left(1+\sum_{\mu=1}^{\infty} K^{-\mu}\psi_\mu(\alpha)\right).$$

Put
$$b_\nu(\alpha)=\frac{(-1)^{\nu-1}}{\nu}\left(\frac{\alpha}{u(\alpha)}\right)^\nu+\frac{\alpha^\nu}{\nu!}f^{(\nu)}(u(\alpha)),$$

$$\exp\{\xi\sum_{\nu=1}^{\infty} b_{\nu+2}(\alpha)\zeta^\nu\}\equiv\sum_{\mu=0}^{\infty} h_\mu(\xi,\alpha)\zeta^\mu;\ h_\mu(\xi,\alpha)=\sum_{\nu=0}^{\mu} a_{\mu\nu}(\alpha)\,\xi^\nu,$$

then

$$\psi_\mu(\alpha)\equiv\sum_{\nu=0}^{2\mu}\frac{\Gamma\left(\mu+\nu+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)}\,a_{2\mu,\nu}(\alpha)\,(2u'(\alpha))^{\nu+\mu}.$$

3. The following additional remarks might widen the field of applicability of Theorem 1. Throughout the previous proof, α was considered a fixed parameter. Now it is clear that nothing will change in the proof if we assume that α is not fixed but depends on k, of course subject to the condition $|\alpha|\leqq\varrho_0<\varrho$. In particular, Theorem 1 remains true if α tends to 0 as $k\to\infty$. For example, the formula of PLANCHEREL and ROTACH for Hermite polynomials holds uniformly if $\varphi\to\infty$ as $k\to\infty$.

If we are interested in the asymptotic behaviour of A_k when k tends less rapidly to ∞ than K, we have to assume that σ is not a constant but tends to 0 as $k\to\infty$. In that case it is preferable to put $\tau=0$, $\sigma=\sigma(k)=\frac{k}{K}$ so that $u(t)=u(\sigma,t)$ depends on k. Again, Theorem 1 remains true, with the additional remark that the $\psi_\mu(\alpha)$ now depend on k and tend to 0 as $k\to\infty$. As a matter of fact, we can always put $\psi_\mu(\alpha)=\frac{k}{K}\chi_\mu(\alpha)$ as readily seen from (35).

Finally let us consider a particular case when the function $f(z)$ itself depends on k. We assume that the coefficients c_ν have an asymptotic development of the form[6])

$$c_\nu=C_\nu\left(1+\sum_{\mu=1}^{m-1}\binom{\nu+1}{\mu}E_\mu K^{-\mu}+\delta_\nu E_m\binom{\nu+1}{m}K^{-m}\left(1+\frac{a}{K}\right)^\nu\right) \tag{38}$$

where C_μ, E_μ, $a\geqq 0$ are constants (not depending on k), and $|\delta_\nu|<1$. Let m be fixed and put $M=\max(|E_\mu|^{\frac{1}{\mu}})$, $\mu=1,\ldots,m$, so that

---

[6]) This case is important from the point of view of partitions to which I have referred above.

$\left|\binom{\nu+1}{\mu} E_\mu\right| \leq \binom{\nu+1}{\mu} M^\mu =$ the coefficient of $K^{-\mu}$ in $\left(1+\frac{M}{K}\right)^{\nu+1}$. Then if $r_1+\ldots+r_i=\nu$,

$$(39)\quad c_{r_1}\ldots c_{r_i} = C_{r_1}\ldots C_{r_i}\prod_{j=1}^{i}\left(1+\sum_{\mu=1}^{m-1}\binom{r_j+1}{\mu}E_\mu K^{-\mu}+\delta E_m\binom{r_j+1}{m}K^{-m}\left(1+\frac{a}{K}\right)^{r_j}\right)$$
$$= C_{r_1}\ldots C_{r_i}\left(1+\sum_{\mu=1}^{m-1}E(r_1,\ldots,r_i;\mu)K^{-\mu}+\delta E(r_1,\ldots,r_i;m)K^{-m}\left(1+\frac{a}{K}\right)^{\nu}\right)$$

where $|\delta|<1$ and

$$(40)\quad |E(r_1,\ldots,r_i;\mu)| < \binom{\nu+i}{\mu}M^\mu \leq \binom{2\nu}{\mu}M^\mu = O(\nu^\mu).$$

Let us write $U(t)$, $V(t)$ for the u, v-functions belonging to $F(z) = \sum_{\nu=1}^{\infty}\frac{1}{\nu}C_\nu z^\nu$ and $\bar\varrho$ for the radius of $V(t)$, then $\varrho \geq \bar\varrho\left(1+\frac{a}{K}\right)^{-1}$ by (19), (39) and (40), hence $\varphi(\sigma,t)$ is convergent for $|t|\leq\varrho_0<\varrho$ if k is large. Also the expansion (36) remains valid as seen by putting (39) into (35) and noticing that the power series belonging to a fixed $k^{-\mu}$, $\mu=1,\ldots,m$, is absolutely convergent by (40).

Theorem 2. *If in Theorem* 1, *the coefficients c_ν have an asymptotic development* (38) *and we put* $F(z)=\sum_{\nu=1}^{\infty}\frac{1}{\nu}C_\nu z^\nu$, $U=t\left(\sigma+\sum_{\nu=1}^{\infty}C_\nu U^\nu\right)$, $V=t\left(\sigma+\sum_{\nu=1}^{\infty}|C_\nu|V^\nu\right)$ *then*

$$\log A_k = k\log K - K\int_0^\alpha\frac{U(t)-\sigma t}{t^2}\,dt+\sum_{\mu=0}^{m}k^{-\mu}\psi_\mu(\alpha)+O(k^{-m-1})$$

uniformly for $|\alpha|\leq\varrho_0<\bar\varrho$ *if* $\bar\varrho$ *is the radius of* $V(t)$. *Also*

$$\psi_0(\alpha)=(\sigma\tau-1)\log\frac{U(\alpha)}{\sigma\alpha}+\frac{1}{2}\log\left(\frac{1}{\sigma}U'(\alpha)\right)$$

if $E_1=0$ *in* (38).

(Received September 21, 1949.)

Remarks on power series.

By G. PÓLYA in Stanford (California).

The following three sections are independent of each other. Each section begins with the statement of a theorem and ends with its short proof. These theorems throw sidelights on various aspects of the theory of power series to which the author devoted much of his work almost since the time that he had the privilege to study under the guidance of Professor LEOPOLD FEJÉR and to make his first contacts with Professor FREDERICK RIESZ.

1. *Assume that $f(z)=e^{-cz}f_1(z)$ where $c\geqq 0$ and $f_1(z)$ is an entire function of genus 0 having positive zeros only. Let γ be the first zero of $f(z)$ and put*

$$-zf'(z)/f(z)=s_1z+s_2z^2+s_3z^3+\dots \tag{1}$$

$$1/f(z)=t_0+t_1z+t_2z^2+t_3z^3+\dots \tag{2}$$

Then s_n/s_{n+1} decrease; and t_n/t_{n+1} increases monotonically so that

$$\frac{t_0}{t_1}\leqq\frac{t_1}{t_2}\leqq\frac{t_2}{t_3}\leqq\dots\leqq\gamma\leqq\dots\leqq\frac{s_2}{s_3}\leqq\frac{s_1}{s_2}. \tag{3}$$

The term "entire function of genus 0" denotes a function of the form

$$f_1(z)=Cz^m\left(1-\frac{z}{\gamma_1}\right)\left(1-\frac{z}{\gamma_2}\right)\left(1-\frac{z}{\gamma_3}\right)\dots \tag{4}$$

$C, m, \gamma_1, \gamma_2, \gamma_3, \dots$ are constants, m an integer, $m\geqq 0$, $0<|\gamma_1|\leqq|\gamma_2|\leqq\dots$, $\Sigma|\gamma_n|^{-1}$ convergent; γ_1 is called the first zero of $f_1(z)$. We do not exclude the case in which the sequence $\gamma_1, \gamma_2, \dots$ is finite or even empty so that, in very special cases, $f_1(z)$ can turn out a polynomial or even a constant. If the γ_n are all positive and $c\geqq 0$, $f(z)=e^{-cz}f_1(z)$ represents the most general function that can be a limit of polynomials with only positive zeros.[1])

---

[1]) G. PÓLYA, Über Annäherung durch Polynome mit lauter reellen Wurzeln, *Rendiconti del Circolo Mat. Palermo,* **36** (1913), pp. 279—295. To the theory of these functions, started by LAGUERRE and developed by I. SCHUR and the author, I. J. SCHOENBERG added recently an important new chapter; see I. J. SCHOENBERG, On totally positive functions, Laplace integrals and entire functions of the Laguerre—Pólya—Schur type, *Proceedings National Academy of Sciences U. S. A.,* **33** (1947), pp. 11—17; On Pólya frequency functions. II: Variation-diminishing integral operators of the convolution type, *these Acta,* **12 B** (1950), pp. 97—106. Prof. SCHOENBERG communicates me a more general theorem which he found some time ago and from which the present result easily follows.

The hypothesis of our theorem requires that $C \neq 0$, $m = 0$ and the sequence $\gamma_1, \gamma_2, \ldots$ has at least one term $\gamma_1 = \gamma$. To the conclusion of our theorem we could give a sharper, but heavier, form by listing the cases of equality in (3) which are few and trivial and will be completely cleared up by the proof.

From the hypothesis of our theorem it follows also that

$$\lim_{n\to\infty} \frac{t_n}{t_{n+1}} = \gamma = \lim_{n\to\infty} \frac{s_n}{s_{n+1}}. \tag{5}$$

These relations are not new. They are due essentially to DANIEL BERNOULLI[2]) whose classical method consists precisely in approximating the minimum root of an equation (as $f(z)=0$) by the ratio of successive coefficients (as s_n/s_{n+1} or t_n/t_{n+1}) of an appropriate power series (as (1) or (2)). Our theorem brings into Bernoulli's method a twofold specialization. First, it chooses a narrow, although quite important, class of functions. Second, it chooses the series (1) and (2). At the price of these specializations, it obtains (3) that is, precise lower and upper bounds for γ at each step of the approximation. Notice that (5) in itself yields no definite estimate for the error of approximation.

Example: Let

$$f(z) = J_0(2z^{\frac{1}{2}}) = 1 - z + \frac{z^2}{4} - \frac{z^3}{36} + \frac{z^4}{576} - \ldots$$

Using just the coefficients displayed, we find

$$\frac{t_3}{t_4} = \frac{304}{211} < \gamma < \frac{48}{33} = \frac{s_3}{s_4}$$

and so for the first root of $J_0(z)$

$$2.4006 < 2\gamma^{\frac{1}{2}} < 2.4121.$$

Proof. Since $f_1(z)$ is of genus 0, s_n, defined by (1), is, for $n \geqq 2$, the sum of the n-th powers of the reciprocal zeros of $f(z)$. Therefore. as well known[3])

$$s_{n+1}^2 \leqq s_n s_{n+2}.$$

For $n \geqq 2$ equality can be attained only if γ is the only, possibly multiple, zero of $f(z)$. Equality for $n=1$ requires, furthermore, $c=0$.

We start discussing t_n.

Lemma. *Assume that $a_n > 0$ for $n \geqq 0$, that*

$$a_n^2 \geqq a_{n-1} a_{n+1} \tag{6}$$

[2]) See L. EULER, *Introductio in Analysin Infinitorum,* Opera Omnia, ser. 1, vol. 8, p. 339.

[3]) See *e. g.* G. H. HARDY, J. E. LITTLEWOOD and G. PÓLYA, *Inequalities* (Cambridge, 1934), p. 28., theorem 18.

for $n \geqq 1$ and that $p > 0$. Put

$$(1-pz)^{-1} \sum_{0}^{\infty} a_n z^n = \sum_{0}^{\infty} A_n z^n.$$

Then, for $n = 1, 2, 3, \ldots$,

$$A_n^2 > A_{n-1} A_{n+1}.$$

Changing $a_n p^{-n}$, $A_n p^{-n}$ and pz into a_n, A_n and z, respectively, we reduce the theorem to the particular case where $p = 1$, which we assume. Then

$$A_n = a_0 + a_1 + \ldots + a_n,$$

$$A_n^2 - A_{n-1} A_{n+1} = A_n^2 - (A_n - a_n)(A_n + a_{n+1}) = A_n a_n - A_n a_{n+1} + a_n a_{n+1} =$$
$$= a_0 a_n + a_1 a_n + \ldots + a_n a_n - a_0 a_{n+1} - a_1 a_{n+1} - \ldots - a_{n-1} a_{n+1} \geqq a_0 a_n > 0,$$

for it follows from the hypothesis (6) that, for $1 \leqq k \leqq n$,

$$\frac{a_0}{a_1} \leqq \frac{a_1}{a_2} \leqq \ldots \leqq \frac{a_{k-1}}{a_k} \leqq \ldots \leqq \frac{a_{n-1}}{a_n} \leqq \frac{a_n}{a_{n+1}},$$

$$a_k a_n \geqq a_{k-1} a_{n+1}.$$

Now, t_n is defined by (2). If $f(z)$ is a polynomial of the first degree, $t_n^2 = t_{n-1} t_{n+1}$. By repeated application of the Lemma, we derive hence that

$$t_n^2 > t_{n-1} t_{n+1} \tag{7}$$

for polynomials of degree $\geqq 2$. By a limiting process, we pass, not to $f(z)$, but to $f(z)/(z-\gamma)$, which is also a limit of polynomials with only positive roots. If $f(z)/(z-\gamma)$ is not a constant, the coefficients in the expansion of its reciprocal in powers of z are all positive and satisfy, if not (7), an inequality which we obtain from (7) by substituting $\geqq$ for $>$. Now, by another application of the Lemma, we obtain (7) unweakened. In short, (7) holds unless $f(z)$ is a polynomial of the first degree.

2. *Let*

$$a_0 + a_1 z + a_2 z^2 + \ldots + a_n z^n + \ldots \tag{1}$$

be a convergent power series but not a polynomial. Then there exists an infinite sequence $\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots$, where $\varepsilon_n = 1$ or -1 for $n \geqq 0$, such that the function represented by

$$\varepsilon_0 a_0 + \varepsilon_1 a_1 z + \varepsilon_2 a_2 z^2 + \ldots + \varepsilon_n a_n z^n + \ldots$$

satisfies no algebraic differential equation.

This theorem is similar to a theorem stated by FATOU which I proved first[4]).

Proof. The theorem is a corollary of an important, almost forgotten, theorem of GRONWALL[5]).

[4]) A. HURWITZ and G. PÓLYA, Zwei Beweise eines von Herrn Fatou vermuteten Satzes, *Acta Math.*, **40** (1916), pp. 179–183.

[5]) H. GRÖNWALL, Sur les fonctions qui ne satisfont à aucune équation différentielle algébrique, *Öfversigt af Kongl. Vetenskaps-Akademiens Förhandlingar,* Stockholm, 1898, p. 387–395.

Since the series (1) is not a polynomial, there exists a sequence of positive integers $l_1, l_2, l_3, \ldots$ such that

$$l_{m+1} > m l_m, \quad a_{l_m} \neq 0,$$

for $m = 1, 2, 3, \ldots$ We set $\varepsilon_n = 1$ if $n = l_1$ or l_2 or $l_3, \ldots$, $\varepsilon_n = -1$ if n is different from all terms of the sequence $l_1, l_2, l_3, \ldots$, and

$$f(z) = \sum_0^\infty a_n z^n, \quad g(z) = \sum_{n=0}^\infty \varepsilon_n a_n z^n, \quad h(z) = 2 \sum_{m=1}^\infty a_{l_m} z^{l_m}.$$

Then

$$(2) \qquad f(z) + g(z) = h(z).$$

Now by GRONWALL's result, $h(z)$ cannot satisfy any algebraic differential equation. Therefore, at least one of the two functions, $f(z)$ and $g(z)$, cannot satisfy any algebraic differential equation, since, in the opposite case, their sum $h(z)$ would also satisfy one[6]).

3. *A power series which satisfies an algebraic equation formally is necessarily convergent* (and so it must satisfy that equation actually).

The term "formal" must be accurately explained. We are given a fixed positive integer q. A formal power series in $z^{1/q}$ (abbreviated in the following as f. p. s.) is defined by an infinite sequence of complex numbers a_μ,

$$(1) \qquad \ldots, a_{-3}, a_{-2}, a_{-1}, a_0, a_1, a_2, \ldots$$

which, however, begins with an infinity of zeros. That is, there is a k such that

$$(2) \qquad a_\mu = 0 \text{ for } \mu < k.$$

The f. p. s. defined by the sequence (1) subject to (2) may be written in either form

$$\sum a_\mu z^{\mu/q}, \quad \sum_{\mu=k}^\infty a_\mu z^{\mu/q}.$$

A f. p. s. may be written as a finite sum if only a finite number of the a_μ differ from 0, or even as the monomial a_0 if $a_\mu = 0$ for $|\mu| > 0$. Observe that the integer k arising in (2) may be different for different f. p. s. Equality, addition, multiplication and differentiation of f. p. s. are so defined as suggested by the particular case in which the series are convergent:

$$(3) \qquad \Sigma a_\mu x^{\mu/q} = \Sigma b_\mu x^{\mu/q} \text{ means } a_\mu = b_\mu \text{ for } \mu = 0, \pm 1, \pm 2, \ldots$$

$$(4) \qquad \Sigma a_\mu x^{\mu/q} + \Sigma b_\mu x^{\mu/q} = \Sigma (a_\mu + b_\mu) x^{\mu/q},$$

$$(5) \qquad \Sigma a_\lambda x^{\lambda/q} \, \Sigma b_\mu x^{\mu/q} = \Sigma c_\nu x^{\nu/q}$$

where

$$(6) \qquad c_\nu = \sum_{\lambda=-\infty}^{\infty} a_\lambda b_{\nu-\lambda}$$

[6]) See E. H. MOORE, Concerning transcendentally transcendental functions, *Math. Annalen*, **48** (1897), pp. 49–74 for general statements about the field formed by the functions satisfying an algebraic differential equation.

(7) $$(\Sigma a_\mu x^{\mu/q})' = \Sigma(\mu/q) a_\mu x^{(\mu-q)/q}.$$

Observe that the sum in (6) is actually finite, by virtue of the condition (2) and of the analogous condition for the b_μ. With these definitions, the f. p. s. form a ring in which the rule holds: *If a product is* 0, *at least one of the factors is necessarily* 0. In fact, assume that neither factor on the left hand side of (5) is 0. Then there exist an l and an m such that

$$a_\lambda = 0 \text{ for } \lambda < l,\ a_l \neq 0;\ b_\mu = 0 \text{ for } \mu < m,\ b_m \neq 0.$$

Then, however, according to (4),

$$c_{l+m} = a_l b_m \neq 0$$

and so the product is different from 0. (Observe that 0 denotes the f. p. s. for which all terms of the sequence (1) are 0.)

Let $P(x, y, y_1, y_2, \ldots, y_n)$ be a polynomial in its $n+2$ variables and w a f. p. s. On the basis of our definitions, the meaning of the equation

(8) $$P(z, w, w', w'', \ldots, w^{(n)}) = 0$$

is completely clear. If (8) holds, we say that w satisfies the differential equation (8) formally — only formally — if w diverges, also actually if w converges. For instance, the f. p. s.

$$w = 1 + 1!\, z + 2!\, z^2 + 3!\, z^3 + \ldots$$

satisfies the differential equation of order 1

$$z^2 w' + (z-1) w + 1 = 0$$

formally and only formally, not actually. It is thinkable that a f. p. s. w satisfies an algebraic equation

(9) $$P(z, w) = 0,$$

which can be regarded as a differential equation of order 0, only formally. Our theorem asserts that this thinkable situation cannot actually arise[7]).

P r o o f. Let n be the degree of the equation (9). Let $w_1, w_2, \ldots, w_n$ denote the convergent Puiseux expansions of the n roots of (9) in the neighborhood of the origin and w a f. p. s. which satisfies (9). We can find a suitable integer q such that $w_1, w_2, \ldots, w_n$ and w are all f. p. s. in $z^{1/q}$. We write (9) in the form

(10) $$A(z)(w-w_1)(w-w_2)\ldots(w-w_n) = 0$$

where $A(z)$ is a polynomial which does not vanish identically. Yet, if the product (10) equals 0, one of its factors must equal 0. That is, w coincides with the convergent Puiseux expansion of one of the roots. Q. E. D.

(Received November 20, 1949.)

[7]) The theorem is an extreme case of results announced elsewhere (G. PÓLYA, Sur les séries entières satisfaisant à une équation différentielle algébrique, *Comptes Rendus Acad. Sci. Paris*, **201** (1935), pp. 444—445), the proof of which, however, has not yet been published.

Méthodes de sommation des séries de Fourier. I.

Par BÉLA SZ.-NAGY à Szeged.

§ 1. Matrices de type F.

Étant donnée une matrice triangulaire infinie

$$\Lambda = (\lambda_{nk}) \qquad (n=0, 1, \ldots; \ k=0, 1, \ldots, n)$$

de nombres réels ou complexes dont $\lambda_{n0} = 1$, attachons à chaque fonction périodique $f(x)$ de période 2π, intégrable[1]) et ayant la série de Fourier

$$f(x) \sim \sum_{k=0}^{\infty} c_k(x), \qquad c_k(x) = a_k \cos kx + b_k \sin kx,$$

les sommes

$$\sigma_n(x) = \sigma_n(\Lambda, f; x) = \sum_{k=0}^{n} \lambda_{nk} c_k(x).$$

Nous dirons que *la matrice Λ est de type F lorsque les sommes $\sigma_n(x)$ tendent vers $f(x)$ en tout point de Lebesgue de $f(x)$ et cela uniformément dans l'intérieur de tout intervalle où $f(x)$ est continue.*

Le prototype d'une telle matrice est celle qui correspond à la méthode classique de FEJÉR de sommer la série de Fourier par le procédé de la moyenne arithmétique: $\Lambda = \left(1 - \frac{k}{n+1}\right)$. Il en est de même pour toutes les matrices de CESÀRO d'ordre $r > 0$, $\Lambda = (A_{n-k}^{(r)} / A_n^{(r)})$ où $A_m^{(r)} = \binom{m+r}{m}$.

En généralisant ces résultats classiques, HILLE et TAMARKIN[2]) ont montré que toute matrice de NÖRLUND $\Lambda = (P_{n-k}/P_n)$, formée à partir d'une suite $\{p_k\}$ de nombres réels ou complexes en posant $P_m = \sum_{h=0}^{m} p_h$, est de type F, à condition qu'il existe une constante C telle que

---

[1]) Dans un intervalle de période et dans le sens de Lebesgue.

[2]) E. HILLE—J. D. TAMARKIN, On the summability of Fourier series. I, *Transactions American Math. Society*, 34 (1932), p. 757 - 783.

$$\left.\begin{array}{ll}(1) & \sum_0^n |p_k| \\ (2) & n|p_n| \\ (3) & \sum_1^n k|p_k-p_{k-1}| \\ (4) & \sum_1^n |P_k|/k\end{array}\right\} < C|P_n|.$$

Ces conditions sont d'ailleurs seulement suffisantes et non pas nécessaires.

Les conditions (1)—(3) sont vérifiées en particulier lorsque $p_n > p_{n+1} > 0$ $(n=0, 1, \ldots)$ (ce qui est le cas par exemple pour les matrices de CESÀRO d'ordre r tel que $0 < r < 1$ où l'on a $p_k = A_k^{(r-1)}$). La condition (4) se trouve alors même nécessaire pour que Λ soit du type F.

En généralisant ce dernier résultat, NIKOLSKY[3]) vient de montrer que, $\Lambda = (\lambda_{nk})$ étant une matrice de nombres réels tels que la suite $\lambda_{n0}, \lambda_{n1}, \ldots, \lambda_{nn}, 0$ est *convexe* ou *concave* pour toute valeur fixée de n, les conditions

$$(5) \qquad \lim_{n\to\infty} \lambda_{nk} = 1 \qquad (k=1, 2, \ldots),$$

$$(6) \qquad |\lambda_{nk}| < C, \qquad \left|\sum_{k=1}^n \frac{\lambda_{nk}}{n-k+1}\right| < C$$

sont suffisantes et nécessaires pour que Λ soit de type F.[4])

Malgré sa généralité, le theorème de NIKOLSKY n'embrasse évidemment pas celui de HILLE et TAMARKIN, les matrices envisagées par ces auteurs ne satisfaisant pas en général à l'hypothèse d'être convexes.

Voici un théorème embrassant, en tant que conditions suffisantes, tous ces théorèmes comme cas particuliers:

Théorème. *Pour que la matrice $\Lambda = (\lambda_{nk})$ soit de type F, il suffit qu'on ait*

$$(A) \qquad \lim_{n\to\infty} \lambda_{nk} = 1 \qquad (k=1, 2, \ldots)$$

et

$$(B) \qquad \sum_{k=0}^{n-1} \left(\sum_{i=n-k}^n \frac{n-k}{i}\right) |\Delta_{nk}^2| < C\,^{5)}$$

où

$$\Delta_{nk}^2 = \lambda_{nk} - 2\lambda_{n,k+1} + \lambda_{n,k+2} \quad (k=0, \ldots, n-1); \quad \lambda_{n0} = 1, \ \lambda_{n,n+1} = 0.$$

---

[3]) S. M. NIKOLSKY, Sur des méthodes linéaires de sommation des séries de Fourier, *Izvestiya Akad. Nauk SSSR*, série math., **12** (1948), p. 259—278 (en russe).

[4]) La nécessité résulte d'une part du fait qu'une suite convergente d'opérations linéaires est bornée dans la sphère unité de l'espace fonctionnel en question, et d'autre part de ce que les fonctions $\cos nx$ et $\nu_n(x) = \sum_{k=1}^n [\cos kx - \cos(2n+2-k)x]/(n-k+1)$ restent inférieures en module à une même constante.

[5]) On désignera par $C, C_1, \ldots$ des constantes ne dépendant pas de n.

La condition (B) est évidemment équivalente à

$$\text{(B')} \qquad \sum_{k=0}^{n-1} (n-k) \log \frac{n}{n-k} \; |\Delta_{nk}^2| < C.$$

Dans le cas où, pour chaque n fixé, les Δ_{nk}^2 sont de même signe, le premier membre de (B) se réduit par deux transformations abéliennes à

$$\left| (\lambda_{n0} - \lambda_{nn}) - \sum_{k=1}^{n} \lambda_{nk}/(n-k+1) \right|;$$

la condition (B) est donc das ce cas une conséquence de celle (6) de NIKOLSKY.

Dans le cas des matrices de NÖRLUND nos conditions prennent la forme

$$\text{(A}_N\text{)} \qquad \lim_{n\to\infty} p_{n-m}/P_n = 0 \qquad (m=0, 1, \ldots),$$

$$\text{(B}_N\text{)} \qquad \sum_{m=1}^{n} \left(\sum_{i=m}^{n} \frac{m}{i} \right) |p_m - p_{m-1}| < C |P_n|.$$

Elles sont moins restrictives que celles de HILLE et TAMARKIN citées plus haut, (A_N) étant une conséquence de (2), (3) et (B_N) une conséquence de (3), (4). En effet, (3) entraîne que

$$\sum_{n-m}^{n} |p_k - p_{k-1}| \leqq \sum_{k=n-m}^{n} \frac{k}{n-m} |p_k - p_{k-1}| < \frac{C}{n-m} |P_n|,$$

donc $|p_{n-m} - p_{n-m-1}|/|P_n| \to 0$ $(n\to\infty)$, et cela pour $m=0, 1, \ldots$. Comme $p_n/P_n \to 0$ par (2), il en résulte de proche en proche que $p_{n-1}/P_n \to 0$, $p_{n-2}/P_n \to 0$ etc. D'autre part, (3) et (4) entraînent que

$$\sum_{m=1}^{i} m |p_m - p_{m-1}| < C |P_i|, \quad \sum_{i=1}^{n} C|P_i|/i < C^2 |P_n|,$$

donc

$$\sum_{i=1}^{n} \left(\sum_{m=1}^{i} m |p_m - p_{m-1}| \right) \Big/ i < C^2 |P_n|,$$

ce qui est équivalent à (B_N).

Pour la matrice $\Lambda = \big((P_n - P_{k-1})/P_n\big)$ où $P_n = \sum_0^n p_k$, $P_{-1}=0$, qu'on appelle aussi matrice "progressive" pour la distinguer de celle "rétrograde" de NÖRLUND, nos conditions prennent la forme

$$\text{(A}_P\text{)} \qquad \lim |P_n| = \infty,$$

$$\text{(B}_P\text{)} \qquad \sum_{k=0}^{n-1} \left(\sum_{i=n-k}^{n} \frac{n-k}{i} \right) |p_k - p_{k+1}| < C |P_n|.$$

Comme

$$\sum_{i=n-k}^{n} \frac{n-k}{i} \leqq k+1,$$

(B_P) est vérifiée en particulier si l'on a (3).[6])

[6]) Le fait que les conditions (A_P), (3) et (2) assurent que la matrice "progressive" Λ soit du type F, est connu; cf. E. HILLE, Summation of Fourier series, *Bulletin American Math. Society,* **38** (1932), p. 505—528, en particulier p. 513.

Observons enfin que si $\lambda_{nk}=\lambda\left(\frac{k}{n+1}\right)$ où $\lambda(u)$ $(0\leqq u\leqq 1)$ est une fonction absolument continue telle que $\lambda(0)=1$, $\lambda(1)=0$, la condition (B) se réduit à ce que l'intégrale

$$\int_0^1 (1-u)\log\frac{1}{1-u}\,|d\lambda'(u)|$$

existe.[7])

§ 2. Quelques conséquences immédiates des hypothèses (A), (B).

Désignons par ν, dans tout ce qui suit, le plus grand entier compris dans $n/2$. Au lieu de λ_{nk}, $\varDelta_{nk}=\lambda_{nk}-\lambda_{n,k+1}$, $\varDelta^2_{nk}=\varDelta_{nk}-\varDelta_{n,k+1}$ nous écrirons, pour abréger, aussi λ_k, $\varDelta_k$ et $\varDelta^2_k$.

Observons d'abord que

$$(n-k)\sum_{n-k}^{n}\frac{1}{i}\geqq(n-k)\frac{k+1}{n}>\begin{cases}\frac{1}{2}(k+1) & \text{pour } 0\leqq k\leqq\nu,\\ \frac{1}{2}(n-k) & \text{pour } \nu\leqq k\leqq n;\end{cases}$$

donc l'hypothèse (B) entraîne

$$\sum_{k=0}^{\nu-1}(k+1)|\varDelta^2_k|<2C,\quad \sum_{k=\nu}^{n-1}(n-k)|\varDelta^2_k|<2C. \tag{7}$$

On obtient par deux transformations abéliennes:

$$\lambda_\nu=\lambda_0-(\lambda_0-\lambda_\nu)=1-\sum_{k=0}^{\nu-1}\varDelta_k=1-\sum_{k=0}^{\nu-1}(k+1)\varDelta^2_k-\nu\varDelta_\nu, \tag{8}$$

ou encore

$$\lambda_\nu=\lambda_\nu-\lambda_{n+1}=\sum_{k=\nu}^{n}\varDelta_k=(n-\nu+1)\varDelta_\nu-\sum_{k=\nu}^{n-1}(n-k)\varDelta^2_k. \tag{9}$$

On déduit de (8) et (9) la relation

$$\lambda_\nu=\frac{n-\nu+1}{n+1}\lambda_\nu+\frac{\nu}{n+1}\lambda_\nu=$$
$$=\frac{n-\nu+1}{n+1}-\frac{n-\nu+1}{n+1}\sum_{k=0}^{\nu-1}(k+1)\varDelta^2_k-\frac{\nu}{n+1}\sum_{k=\nu}^{n-1}(n-k)\varDelta^2_k.$$

Grâce à (7), cela donne

$$|\lambda_\nu|\leqq 1+4C=C_1. \tag{10}$$

---

[7]) Nous avons étudié des procédés de sommation engendrées par de telles fonctions "sommatoires" $\lambda(u)$ (et même par d'autres plus générales) dans un Mémoire précédent: Sur une classe générale de procédés de sommation pour les séries de Fourier, *Hungarica Acta Math.*, **1**, n° 3 (1948), p. 14–52. Nous avons déterminé les constantes de Lebesgue et les constantes d'approximation correspondantes.

Écrivons (8) sous la forme

$$\nu\Delta_\nu = 1 - \lambda_\nu - \sum_{k=0}^{\nu-1}(k+1)\Delta_k^2.$$

Il en résulte, par (7) et (10), que

(11) $$|\nu\Delta_\nu| \leqq 1 + C_1 + 2C = C_2.$$

Montrons enfin que

(12) $$\Sigma_n = \sum_{k=0}^{n-1} |\Delta_{nk}^2| \to 0 \qquad \text{lorsque} \qquad n \to \infty.$$

Pour le voir, commençons par observer que, quel que soit l'entier $r \geqq 1$, on peut déterminer $n_0 = n_0(r)$ de façon qu'on ait

$$\sum_m^n \frac{1}{i} > r \qquad \text{pour tout } n \geqq n_0 \text{ et pour tout } m \leqq r.$$

On a alors pour $n \geqq n_0$:

$$\Sigma_n \leqq \sum_{k=0}^{r-1} |\Delta_k^2| + \frac{1}{r}\sum_{k=r}^{\nu-1}(k+1)|\Delta_k^2| + $$
$$+ \frac{1}{r}\sum_{k=\nu}^{n-r}(n-k)|\Delta_k^2| + \frac{1}{r}\sum_{k=n-r+1}^{n-1}\left(\sum_{i=n-k}^{n}\frac{n-k}{i}\right)|\Delta_k^2|.$$

En vertu de l'hypothèse (B) et de ses conséquences (7), il en vient que

$$\Sigma_n \leqq \sum_{k=0}^{r-1} |\Delta_k^2| + \frac{5C}{r} \qquad (n \geqq n_0).$$

Étant donné $\varepsilon > 0$ arbitraire, on détermine r de façon que $5C/r < \varepsilon/2$; grâce à l'hypothèse (A), on peut déterminer ensuite n_1 ($\geqq n_0$) de façon que pour $n \geqq n_1$ et pour $k = 0, 1, \ldots, r-1$ on ait $|\Delta_{nk}^2| < \varepsilon/2r$. On aura alors $\Sigma_n < \varepsilon$, ce qui prouve (12).

§ 3. Démonstration du théorème.

Partons de la formule

(13) $$\sigma_n(x) - f(x) = \frac{1}{\pi}\int_0^\pi K_n(t)\,\varphi_x(t)\,dt$$

où

$$\varphi_x(t) = f(x+t) + f(x-t) - 2f(x) \quad \text{et} \quad K_n(t) = \frac{1}{2} + \sum_{k=1}^{n} \lambda_{nk}\cos kt.$$

On obtient par deux transformations abéliennes:

$$K_n(t) = \sum_{k=0}^{n-1} M_k(t)\,\Delta_k^2 + M_n(t)\,\Delta_n,$$

ou encore

(14) $$K_n(t) = \sum_{k=0}^{\nu-1} M_k(t)\,\Delta_k^2 + \sum_{k=\nu}^{n-1} N_k(t)\,\Delta_k^2 + M_n(t)\,\Delta_\nu$$

où

$$M_k(t)=\frac{\sin^2(k+1)t/2}{2\sin^2 t/2},\quad N_k(t)=N_{nk}(t)=M_k(t)-M_n(t).$$

Envisageons un intervalle (δ,π), $\delta>0$. Les fonctions $|M_k(t)|$ et $|N_k(t)|$ y restant inférieures à $\omega(\delta)=\sin^{-2}\delta/2$, on a

$$|K_n(t)|\leqq\omega(\delta)\left(\sum_{k=0}^{n-1}|\Delta_k^2|+|\Delta_\nu|\right)\qquad(\delta\leqq t\leqq\pi),$$

d'où il s'ensuit par (11) et (12) que

(15) $$m_n(\delta)=\max_{(\delta,\,\pi)}|K_n(t)|\to 0\qquad(n\to\infty).$$

Montrons que $|K_n(t)|$ admet dans $(0,\pi)$ une majorante décroissante $K_n^*(t)$ telle que

(16) $$\int_0^\pi K_n^*(t)\,dt<C_3.$$

Pour $|M_k(t)|$ on a évidemment la majorante décroissante

$$M_k^*(t)=\begin{cases}\left[(k+1)\dfrac{t}{2}\right]^2\Big/2\left(\dfrac{t}{\pi}\right)^2=\pi^2(k+1)^2/8 & \left(0\leqq t\leqq\dfrac{2}{k+1}\right),\\ 1\Big/2\left(\dfrac{t}{\pi}\right)^2=\pi^2/2t^2 & \left(\dfrac{2}{k+1}\leqq t\leqq\pi\right)\end{cases}$$

et

(17) $$\int_0^\pi M_k^*(t)\,dt\leqq C_4(k+1).$$

D'autre part,

$$|N_k(t)|=\frac{|\cos(n+1)t-\cos(k+1)t|}{2\sin^2 t/2}=\frac{|\sin(n+k+2)t/2\cdot\sin(n-k)t/2|}{\sin^2 t/2}$$

admet la majorante décroissante

$$N_k^*(t)=\begin{cases}\dfrac{(n+k+2)t/2\cdot(n-k)t/2}{(t/\pi)^2}=\pi^2(n+k+2)(n-k) & \left(0\leqq t\leqq\dfrac{2}{n+k+2}\right),\\ \dfrac{(n-k)\,t/2}{(t/\pi)^2}=\pi^2\dfrac{n-k}{2t} & \left(\dfrac{2}{n+k+2}\leqq t\leqq\dfrac{2}{n-k}\right),\\ \dfrac{1}{(t/\pi)^2}=\pi^2\dfrac{1}{t^2} & \left(\dfrac{2}{n-k}\leqq t\leqq\pi\right),\end{cases}$$

et on a, pour $0\leqq k\leqq n$,

$$\int_0^\pi N_k^*(t)\,dt=\frac{\pi^2}{2}(n-k)\left[2+\log\frac{n+k+2}{n-k}\right]-\pi<$$
$$<\frac{\pi^2}{2}(n-k)\left[2+\log 2+\log\frac{n+1}{n-k}\right],$$

donc

(18) $$\int_0^\pi N_k^*(t)\,dt\leqq(n-k)\left(C_5+C_6\sum_{i=n-k}^{n}\frac{1}{i}\right).$$

Pour $|K_n(t)|$ on obtient ainsi la majorante décroissante

$$K_n^*(t)=\sum_{k=0}^{\nu-1} M_k^*(t)\,|\varDelta_k^2|+\sum_{k=\nu}^{n-1} N_k^*(t)\,|\varDelta_k^2|+M_n^*(t)\,|\varDelta_\nu|.$$

Par (17) et (18),

$$\int^{\pi} K_n^*(t)\,dt \leqq C_4\sum_{k=0}^{\nu-1}(k+1)\,|\varDelta_k^2|+\sum_{k=\nu}^{n-1}(n-k)\left(C_5+C_6\sum_{i=n-k}^{n}\frac{1}{i}\right)|\varDelta_k^2|+C_4(\nu+1)\,|\varDelta_\nu|,$$

d'où il résulte par (B) et par ses conséquences (7), (11), l'inégalité (16).

Or les propriétés

$$m_n(\delta)\to 0 \quad (n\to\infty), \qquad \int_0^{\pi}|K_n(t)|\,dt<C_3$$

suffisent à établir la convergence uniforme de $\sigma_n(x)$ vers $f(x)$ dans l'intérieur de tout intervalle (a,b) où $f(x)$ est continue, on n'a qu'à se servir de (13) et de l'évaluation

$$\left|\int_0^{\pi}K_n(t)\,\varphi_x(t)\,dt\right|\leqq \max_{0\leqq t\leqq\delta}|\varphi_x(t)|\int_0^{\pi}|K_n|\,dt+m_n(\delta)\int_0^{\pi}|\varphi_x(t)|\,dt.$$

Pour montrer que $\sigma_n(x)$ tend vers $f(x)$ en tout point de Lebesgue, ou d'une manière plus précise, en tout point x tel que

$$\varPhi_x(t)=\int_0^t|\varphi_x(s)|\,ds=o(t) \quad \text{pour} \quad t\to 0,$$

on fera usage aussi des majorantes décroissantes et de (16). Pour un tel x, on aura $\varPhi_x(t)\leqq\varepsilon t$ dès que $0\leqq t\leqq\delta=\delta(\varepsilon)$ et par conséquent

$$\left|\int_0^{\delta}K_n(t)\,\varphi_x(t)\,dt\right|\leqq\int_0^{\delta}K_n^*(t)\,|\varphi_x(t)|\,dt=[\varPhi_x(t)\,K_n^*(t)]_0^{\delta}+\int_0^{\delta}\varPhi_x(t)\,d[-K_n^*(t)]\leqq$$
$$\leqq[\varepsilon t\,K_n^*(t)]_0^{\delta}+\int_0^{\delta}\varepsilon t\,d[-K_n^*(t)]=\varepsilon\int_0^{\delta}K_n^*(t)\,dt<C_3\varepsilon.$$

D'autre part, pour n assez élevés,

$$\left|\int_{\delta}^{\pi}K_n(t)\,\varphi_x(t)\,dt\right|\leqq m_n(\delta)\int_0^{\pi}|\varphi_x(t)|\,dt<\varepsilon.$$

On a donc

$$\left|\int_0^{\pi}K_n(t)\,\varphi_x(t)\,dt\right|<(1+C_3)\,\varepsilon.$$

Comme ε était arbitraire, cela achève la démonstration du théorème.

Dans une seconde communication on traitera des problèmes analogues pour la sommation des séries conjuguées.

(Reçu le 25 juin 1949)

Sur les nombres de Lipschitz approximatifs.

Par ÁKOS CSÁSZÁR à Budapest.

$F(x)$ étant une fonction réelle finie d'une variable réelle x, on appelle suivant BESICOVITCH [1] nombres de Lipschitz d'ordre α $(0<\alpha<1)$ de $F(x)$ au point x_0 les limites suivantes:

$$\overline{L}_\alpha^+ F(x_0)=\limsup_{x\to x_0+0}\frac{F(x)-F(x_0)}{(x-x_0)^\alpha},\quad \underline{L}_\alpha^+ F(x_0)=\liminf_{x\to x_0+0}\frac{F(x)-F(x_0)}{(x-x_0)^\alpha},$$

$$\overline{L}_\alpha^- F(x_0)=\limsup_{x\to x_0-0}\frac{F(x_0)-F(x)}{(x_0-x)^\alpha},\quad \underline{L}_\alpha^- F(x_0)=\liminf_{x\to x_0-0}\frac{F(x_0)-F(x)}{(x_0-x)^\alpha}.$$

Nous définissons les nombres de Lipschitz approximatifs $\overline{A}_\alpha^+ F(x_0)$, $\underline{A}_\alpha^+ F(x_0)$, $\overline{A}_\alpha^- F(x_0)$, $\underline{A}_\alpha^- F(x_0)$ par les mêmes formules, en y remplaçant la limite supérieure (resp. inférieure) par la limite supérieure (resp. inférieure) approximative[1]).

Dans son mémoire cité, BESICOVITCH démontre le théorème suivant: Pour toute fonction mesurable $F(x)$, l'ensemble des points x pour lesquels on a $\overline{L}_\alpha^- F(x)<\overline{L}_\alpha^+ F(x)<+\infty$, est de mesure nulle. Dans cette note, nous donnerons une démonstration plus simple de ce théorème, valable pour des fonctions arbitraires (mesurables ou non). D'autre part, nous établirons une proposition analogue pour les nombres de Lipschitz approximatifs.

Théorème 1. (Généralisation du théorème de BESICOVITCH.) *Pour toute fonction finie $F(x)$, l'ensemble des points x pour lesquels on a*

$$\overline{L}_\alpha^- F(x)<\overline{L}_\alpha^+ F(x)<+\infty,$$

est de mesure nulle.

Démonstration. On a $\overline{D}^- F(x)=-\infty$ pour tout point où on a $\overline{L}_\alpha^- F(x)<0$, de sorte que cette inégalité ne peut avoir lieu que dans les points d'un ensemble de mesure nulle. Il suffit donc évidemment de montrer

[1]) La limite supérieure approximative à droite d'une fonction $f(x)$ au point x_0 est la borne inférieure des nombres y pour lesquels l'ensemble des valeurs x satisfaisant à l'inégalité $f(x)>y$ a la densité extérieure à droite 0 au point x_0.

que pour $0<p<q<r$ l'ensemble $E_{p,q,r}$ des points x satisfaisant à l'inégalité

$$0 \leqq \bar{L}_\alpha^- F(x) < p < q < \bar{L}_\alpha^+ F(x) < r$$

est de mesure nulle. Désignons par η un nombre positif tel que $p+r\eta^\alpha<q$. Supposons que $|E_{p,q,r}|>0$. En désignant par $E_{p,q,r,n}$ l'ensemble des points $x \in E_{p,q,r}$ pour lesquels $x<t\leqq x+\frac{1}{n}$ entraîne $F(t)-F(x)<r(t-x)^\alpha$ et $x-\frac{1}{n}\leqq t<x$ entraîne $F(x)-F(t)<p(x-t)^\alpha$, on a $E_{p,q,r}=\sum_{n=1}^{\infty} E_{p,q,r,n}$, de sorte que pour une valeur de n on a $|E_{p,q,r,n}|>0$. x_0 soit un point de densité extérieure de $E_{p,q,r,n}$. On a alors en vertu de l'inégalité $L_\alpha^+ F(x_0)>q$ la relation

$$(1) \qquad F(x)-F(x_0)>q(x-x_0)^\alpha$$

pour des valeurs $x>x_0$ arbitrairement voisines de x_0; x_1 soit une de ces valeurs assez voisine de x_0 pour que x_1-x_0 soit moindre que $\frac{1}{n}$ et que la mesure extérieure de la partie de l'ensemble $E_{p,q,r,n}$ située entre x_0 et x_1 dépasse le nombre $\eta(x_1-x_0)$. On peut alors trouver des points de $E_{p,q,r,n}$ entre $x_1-\eta(x_1-x_0)$ et x_1; x_2 soit un tel point. On a donc les inégalités

$$F(x_2)-F(x_0)<p(x_2-x_0)^\alpha \leqq p(x_1-x_0)^\alpha$$

et

$$F(x_1)-F(x_2)<r(x_1-x_2)^\alpha \leqq r\eta^\alpha(x_1-x_0)^\alpha,$$

et par l'addition de ces deux inégalités

$$F(x_1)-F(x_0)<(p+r\eta^\alpha)(x_1-x_0)^\alpha<q(x_1-x_0)^\alpha,$$

ce qui est impossible en vertu de (1).

Théorème 2. *Pour toute fonction finie et mesurable $F(x)$, l'ensemble des valeurs de x pour lesquelles on a*

$$\bar{\Lambda}_\alpha^- F(x) < \bar{\Lambda}_\alpha^+ F(x) < +\infty,$$

est de mesure nulle.

Démonstration. Comme $\bar{\Lambda}_\alpha^- F(x)<0$ entraîne $\bar{D}_{ap}^- F(x)=-\infty$, on a presque partout $\bar{\Lambda}_\alpha^- F(x)\geqq 0$.[2]) Il suffit donc de montrer que, $E_{p,q,r}$ désignant l'ensemble des points x satisfaisant aux inégalités

$$0 \leqq \bar{\Lambda}_\alpha^- F(x) < p < q < \bar{\Lambda}_\alpha^+ F(x) < r,$$

on a $|E_{p,q,r}|=0$ pour $0<p<q<r$.

Considérons à ce but pour $0\leqq t\leqq 1$ la fonction

$$G(t)=p(1-t)^\alpha+rt^\alpha-q.$$

On a $G(0)=p-q<0$, $G(1)=r-q>0$ et une discussion facile montre que $G(t)$ atteind une valeur maximale pour une valeur t_0 située entre 0 et 1, et qu'il croît dans l'intervalle $0\leqq t\leqq t_0$ et décroît dans l'intervalle $t_0\leqq t\leqq 1$.

[2]) Voir p. ex. Saks [2].

On a donc une valeur unique $\lambda = \lambda(p, q, r)$ pour laquelle $G(\lambda) = 0$, $0 < \lambda < 1$ et $G(t)$ est monotone croissant pour $0 \leqq t \leqq \lambda$.

Ceci établi, désignons par $E_{p, q, r, n}$ l'ensemble des points $x \in E_{p, q, r}$ pour lesquels $0 < h \leqq \frac{1}{n}$ entraîne les inégalités

$$|\underset{t}{\mathbf{E}}[F(t) - F(x) > r(t-x)^{\alpha};\ x < t < x+h]| < \frac{\lambda}{3} h,$$

$$|\underset{t}{\mathbf{E}}[F(x) - F(t) > p(x-t)^{\alpha};\ x-h < t < x]| < \frac{\lambda}{3} h.$$

($\underset{t}{\mathbf{E}}[P]$ désigne ici l'ensemble des points t jouissant de la propriété P.)

Désignons enfin par $E_{p, q, r, n, k}$ l'ensemble des points $x \in E_{p, q, r, n}$ dans lesquels la densité extérieure supérieure à droite de l'ensemble

$$\underset{t}{\mathbf{E}}[F(t) - F(x) > q(t-x)^{\alpha};\ x < t]$$

dépasse le nombre $1/k$. On a évidemment

$$E_{p, q, r} = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} E_{p, q, r, n, k},$$

de sorte qu'on n'a qu'à démontrer que $|E_{p, q, r, n, k}| = 0$.

Supposons donc que $A = E_{p_0, q_0, r_0, n_0, k_0}$ soit de mesure extérieure positive pour certaines valeurs p_0, q_0, r_0, n_0, k_0. Désignons alors par x_0 un point de densité extérieure de A appartenant à A. On trouve alors des nombres h arbitrairement petits tels qu'on ait

$$(2) \qquad |\underset{t}{\mathbf{E}}[F(t) - F(x_0) > q(t-x_0)^{\alpha};\ x_0 < t < x_0+h]| > \frac{1}{k} h.$$

Soit h_0 un de ces nombres assez petit pour que les inégalités

$$(3) \qquad 0 < h_0 < \frac{1}{n} \quad \text{et} \quad |A \cdot [x_0, x_0+h_0)| > \left(1 - \frac{1}{2k}\right) h_0$$

soient satisfaites. $F(x)$ étant mesurable, l'ensemble figurant dans le membre gauche de (2) est mesurable, de sorte qu'on a

$$(4) \qquad |\underset{t}{\mathbf{E}}[F(t) - F(x_0) \leqq q(t-x_0)^{\alpha};\ x_0 < t < x_0+h_0]| < \left(1 - \frac{1}{k}\right) h_0.$$

En vertu de (3) et (4), un point x_1 existe tel que $x_0 < x_1 < x_0 + h_0$, $x_1 \in A$ et

$$(5) \qquad F(x_1) - F(x_0) > q(x_1 - x_0)^{\alpha}.$$

x_0 et x_1 appartenant à A, on a en vertu de $0 < x_1 - x_0 < \frac{1}{n}$ les inégalités

$$(6) \qquad |\underset{t}{\mathbf{E}}[F(t) - F(x_0) > r(t-x_0)^{\alpha};\ x_0 < t < x_1]| < \frac{\lambda}{3}(x_1 - x_0),$$

$$(7) \qquad |\underset{t}{\mathbf{E}}[F(x_1) - F(t) > p(x_1-t)^{\alpha};\ x_0 < t < x_1]| < \frac{\lambda}{3}(x_1 - x_0).$$

(6) et (7) entraînent qu'un point x_2 existe tel que

$$x_0 < x_2 < x_0 + \lambda(x_1 - x_0)$$

et

$$F(x_2) - F(x_0) \leqq r(x_2 - x_0)^\alpha, \quad F(x_1) - F(x_2) \leqq p(x_1 - x_2)^\alpha.$$

En écrivant $x_2 - x_0 = \mu(x_1 - x_0)$, on a alors

$$F(x_1) - F(x_0) \leqq p(x_1 - x_2)^\alpha + r(x_2 - x_0)^\alpha = [p(1-\mu)^\alpha + r\mu^\alpha](x_1 - x_0)^\alpha.$$

Mais $0 < \mu < \lambda$ entraîne en vertu de ce que nous venons d'établir sur la monotonité de la fonction $G(t)$ que $G(\mu) < 0$, c'est-à-dire que $p(1-\mu)^\alpha + r\mu < q$, ce qui donne enfin

$$F(x_1) - F(x_0) < q(x_1 - x_0)^\alpha,$$

ce qui contredit à l'inégalité (5). C. Q. F. D.

Corollaire. *$F(x)$ étant une fonction finie et mesurable satisfaisant aux points d'un ensemble E aux inégalités $\bar{\Lambda}_\alpha^+ F(x) < +\infty$, $\bar{\Lambda}_\alpha^- F(x) < +\infty$, on a $\bar{\Lambda}_\alpha^+ F(x) = \bar{\Lambda}_\alpha^- F(x)$ presque partout dans E.*

Ouvrages cités.

[1] A. S. Besicovitch, On Lipschitz numbers, *Math. Zeitschrift*, **30** (1929), p. 514–519.
[2] S. Saks, *Theory of the integral* (Warszawa–Lwów, 1937), p. 295.

(Reçu le 14 septembre 1949)

On the geometry of conformal mapping.

By ALFRÉD RÉNYI in Budapest.

Introduction.

Let us denote by S the class of analytic functions

$$f(z)=z+a_2z^2+\ldots+a_nz^n+\ldots \tag{1}$$

which are regular and schlicht in the circle $|z|<1$. Let us denote by $D(r)$ the domain of the w-plane onto which the circle $|z|<r$ $(r<1)$ is mapped by the function $w=f(z)$, and by $C(r)$ the boundary of $D(r)$. Let $A(r)$ denote the area of $D(r)$ and $L(r)$ the length of the curve $C(r)$. We put $z=re^{i\varphi}$ and denote by $s=s(r,\varphi)$ the length along $C(r)$ from the point $w=f(r)$ to the point $f(re^{i\varphi})$ in the positive direction. We have evidently

$$\frac{ds}{d\varphi}=r|f'(z)|. \tag{2}$$

Let us put $\arg f'(z)=\chi$ and $\psi=\chi+\varphi+\frac{\pi}{2}$; clearly ψ denotes the angle between the tangent to $C(r)$ in the point $f(re^{i\varphi})$ and the real axis of the w-plane. Let us denote by $R=R(r,\varphi)$ the radius of curvature of $C(r)$ in the point $f(re^{i\varphi})$ and let us put $\gamma=\gamma(r,\varphi)=\frac{1}{R(r,\varphi)}$; it follows

$$\gamma=\frac{d\psi}{ds}=\frac{1+R\left(\frac{zf''(z)}{f'(z)}\right)}{r|f'(z)|}. \tag{3}$$

Here and in what follows we denote by $R(\zeta)$ the real part, and by $I(\zeta)$ the imaginary part of the complex number ζ. We denote by $S(f)=S(f(z))$ the invariant of SCHWARZ[1])

$$S(f)=\frac{f'''(z)}{f'(z)}-\frac{3}{2}\left(\frac{f''(z)}{f'(z)}\right)^2. \tag{4}$$

---

[1]) The invariant of SCHWARZ is the differential form of least order which remains invariant with respect to every linear transformation effected on $f(z)$; cf. H. A. SCHWARZ, *Gesammelte Math. Abhandlungen.* II (Berlin, 1890), pp. 351—355.

It is known, that for any $f(z)$, belonging to S, $C(r)$ is convex for $0 < r < r_c = 2 - \sqrt{3} = 0.26\ldots$, and star-like with respect to the point $w=0$ for $r < r_s = \tanh \frac{\pi}{4} = 0.65\ldots$[2]). In Part I we shall investigate in detail the form of $C(r)$ for $r < r_c$; it is evident that when r decreases, the form of $C(r)$ approaches more and more the form of a circle, our aim is to express this fact in a precise manner. For this purpose we have to introduce some quantity measuring the degree of dissemblance between $C(r)$ and a circle; for this quantity we choose the total variation of γ along $C(r)$, i. e. we put

$$\delta(r) = \int_{C(r)} |d\gamma| = \int_0^{2\pi} \left| \frac{d\gamma}{d\varphi} \right| d\varphi. \tag{5}$$ [3])

The following theorem will be proved:

Theorem 1. *For any function belonging to S we have*

$$\delta(r) < \frac{12\pi r(1+r)}{(1-r)^3}. \tag{6}$$

As a consequence of Theorem 1 (Corollary II) we shall prove that $C(r)$ is contained between two circles with radii $r - O(r^3)$ and $r + O(r^3)$. This is an improvement compared with the distortion theorem[4]), from which it follows only that $C(r)$ is contained between two circles with radii $r \pm O(r^2)$. (Here and in what follows we denote by $O(r^2)$, $O(r^3)$ etc. quantities which are bounded uniformly (i. e. independently of r as well as of $f(z)$) when divided by r^2, r^3, etc.) To prove the mentioned result we need the following

Lemma 1. *For any $f(z)$ belonging to S we have*

$$L(r) = 2\pi r + O(r^3). \tag{8}$$

Clearly Lemma 1 can be expressed also by stating that

$$\left(\frac{d^2 L}{dr^2} \right)_{r=0} = 0. \tag{9}$$

---

[2]) The radius of convexity has been determined by R. Nevanlinna, Über die schlichte Abbildungen des Einheitskreises, *Översigt av Finska Vet. Soc. Forhandlingen*, **62** (1920), pp. 1—14; the exact radius of starlikeness has been found, after long series of trials, by H. Grunsky, Zwei Bemerkungen zur konformen Abbildung, *Jahresbericht der Deutschen Math. Vereinigung*, **5** (1933), pp. 140—143.

[3]) The use of this quantity has been kindly suggested to me by Dr. István Fáry. It must be added, that the quantity defined by (5) gives a measure of the dissemblance of a curve from the circle only if the knowledge of the size of the curve (f. e. its length is presupposed; an absolute measure of dissemblance is furnished by the product of (5) with the length of the curve.

[4]) The distortion theorem asserts that $C(r)$ is contained between the two concentric circles with centre at the origin having the radii $\frac{r}{(1+r)^2}$ and $\frac{r}{(1-r)^2}$.

It may be mentioned that (8) is by no means evident, as from the distortion theorem[6]) applied to the formula

$$L(r)=r\int_0^{2\pi}|f'(z)|\,d\varphi \tag{10}$$

it follows at the first sight only $L(r)=2\pi r+O(r^2)$.

In Part II we investigate the form of $C(r)$ for $r_c<r<r_s$. We define $K(r)$, the set of those (interior) points of $D(r)$, with respect to which $C(r)$ is star-like; we shall call $K(r)$ the star-kernel of $D(r)$[5]). According to the theorems mentioned above, and taking into account that a convex domain is star-like with respect to every of its interior points, it follows that $K(r)=D(r)$ for $r\leqq r_c$ and $K(r)$ not void for $r\leqq r_s$. The question arises what can be said regarding the size of $K(r)$ for $r_c<r<r_s$. Theorem 2 is a first attempt to answer this question.

In the present paper we do not consider the range of values $r_s<r<1$, we refer only to the interesting results obtained by GOLUSIN[6]).

Part I.

We shall need the following

Lemma 2. *For any function $f(z)=z+a_2z^2+a_3z^3+\dots$ belonging to S we have $|a_2^2-a_3|\leqq 1$. This inequality is best possible as equality stands for* $f(z)=\dfrac{z}{(1-z)^2}$. This lemma has been proved by GOLUSIN[7]) and SCHIFFER[8]).

Using Lemma 2 we obtain the following estimation of the invariant of SCHWARZ:

Lemma 3. *For any $f(z)$ belonging to S we have*

$$|S(f)|=\left|\frac{f'''(z)}{f'(z)}-\frac{3}{2}\left(\frac{f''(z)}{f'(z)}\right)^2\right|\leqq\frac{6}{(1-r^2)^2}. \tag{15}$$

This is a "best possible" result as for $f(z)=\dfrac{z}{(1-z)^2}$ and $z=r$ we have equality in (15).

---

5) It is easy to see that $K(r)$ is a convex domain. This has been mentioned first by Thekla LUKÁCS; cf. G. PÓLYA and G. SZEGŐ, *Aufgaben und Lehrsätze aus der Analysis.* I. (Berlin, 1925), p. 277.

6) For the mentioned results and for further literature we refer to the excellent survey article of G. M. GOLUSIN, Interior problems of the theory of schlicht functions, *Uspekhi Mat. Nauk*, **6** (1939), pp. 26—89.

7) G. M. GOLUSIN, Einige Koeffizientenabschätzungen für schlichte Funktionen, *Mat. Sbornik*, **3** (1938), pp. 321—330.

8) M. SCHIFFER, Sur un problème d'extremum de la représentation conforme, *Bulletin de la Société Math. de France*, **66** (1938), pp. 48—55.

To prove Lemma 3 let us introduce the function

$$h(\zeta)=\frac{f\left(\dfrac{\zeta+z}{1+\bar{z}\zeta}\right)-f(z)}{f'(z)(1-r^2)}. \tag{16}$$

A simple calculation gives

$$(1-r^2)^2 S(f)=h'''(0)-\frac{3}{2}(h''(0))^2. \tag{17}$$

As $h(\zeta)$ belongs evidently to S, putting $(\zeta)=\zeta+c_2\zeta^2+c_3\zeta^3+\dots$ and applying Lemma 2, we have $(1-r^2)^2|S(f)|=6|c_3-c_2^2|\leqq 6$ which proves Lemma 3. Lemma 3 has been proved recently in another way (without using Lemma 1) by NEHARI[9]).

Let us calculate now the variation of γ along $C(r)$. We have by some calculations

$$\frac{d\gamma}{d\varphi}=\frac{I[z^2S(f)]}{r|f'(z)|}. \tag{18}$$

Using the distortion theorem, according to which

$$\frac{1-r}{(1+r)^3}\leqq|f'(z)|\leqq\frac{1+r}{(1-r)^3} \tag{19}$$

and using Lemma 3, Theorem 1 follows immediately.

Before considering the consequences of Theorem 1 we prove Lemma 1. We start by the decomposition

$$\begin{aligned}L(r)=r\int_0^{2\pi} f'(z)e^{-i\chi}d\varphi=r\int_0^{2\pi} f'(z)\,d\varphi+r\int_0^{2\pi}(e^{-i\chi}-1)\,d\varphi+\\ +r\int_0^{2\pi}(f'(z)-1)(e^{-i\chi}-1)\,d\varphi.\end{aligned} \tag{20}$$

Evidently

$$r\int_0^{2\pi} f'(z)\,d\varphi=2\pi r\frac{1}{2\pi i}\int\frac{f'(z)\,dz}{z}=2\pi r, \tag{21}$$

further

$$r\int_0^{2\pi}(e^{-i\chi}-1)\,d\varphi=-ir\int_0^{2\pi}\chi\,d\varphi+r\int_0^{2\pi}(e^{-i\chi}-1+i\chi)\,d\varphi. \tag{22}$$

As $\log f'(z)$ is regular in $|z|<1$, $\chi=I(\log f'(z))$ is a harmonic function, and thus $\int_0^{2\pi}\chi d\varphi=\chi(0)=0$; using the elementary inequality $|e^{-ix}+ix-1|=O(x^2)$ and the rotation theorem:

---

[9]) Z. NEHARI, The Schwarzian derivative and schlicht functions, *Bulletin of the American Math. Society*, **55** (1949), pp. 545—551.

(23) $$|\chi| \leqq 2 \log \frac{1+r}{1-r}$$

we obtain

(24) $$r \int_0^{2\pi} (e^{-i\chi} - 1)\, d\varphi = O(r^3).$$

As regards the third term of (20), we have by (19) and (23)

(25) $$r \left| \int_0^{2\pi} (f'(z) - 1)(e^{-i\chi} - 1)\, d\varphi \right| \leqq r \int_0^{2\pi} |f'(z) - 1|\, |\chi|\, d\varphi = O(r^3).$$

Thus, using (20), (21), (24) and (25), it follows

(26) $$L(r) = 2\pi r + O(r^3)$$

which is Lemma 1.

As an immediate consequence of Lemma 1 we mention that the isoperimetric deficiency of $C(r)$ is $O(r^4)$. As a matter of fact, we have by a well known formula

$$A(r) = \pi r^2 + \sum_{n=2}^{\infty} n^2 |a_n|^2 r^{2n}$$

from which, combined with (26) it follows

(27) $$L^2(r) - 4\pi A(r) = O(r^4).$$

Now let us consider some consequences of Theorem 1. We start by the formula

(28) $$L(r) = \int_0^{2\pi} R\, d\psi.$$

Let us denote by R_0 the mean value of R on $C(r)$, i. e. we put

(29) $$R_0 = \frac{1}{2\pi} \int_0^{2\pi} R\, d\psi = \frac{L(r)}{2\pi}.$$

For any value of R we have evidently

(30) $$\left| \frac{1}{R} - \frac{1}{R_0} \right| \leqq \delta(r)$$

and thus, for sufficiently small r,

(31) $$\frac{R_0}{1 + R_0 \delta(r)} \leqq R \leqq \frac{R_0}{1 - R_0 \delta(r)}.$$

According to (26) we have $R_0 = r + O(r^3)$ and by Theorem 1 it follows $\delta(r) = O(r)$, thus we have from (31):

Corollary I. *If $R(r,\varphi)$ denotes the radius of curvature of $C(r)$ in the point $f(re^{i\varphi})$ we have*

$$R(r,\varphi)=r+O(r^3) \tag{32}$$

uniformly in φ, r and $f(z)\in S$.

Let us denote by R_M and r_m the maximal resp. minimal value of R on $C(r)$. According to a theorem of BLASCHKE[10]), if the convex curves C_1 and C_2 have a common tangent in one point and the radius of curvature of C_1 exceeds the radius of curvature of C_2 in points with parallel directions, it follows that C_2 is contained in C_1. Thus $C(r)$ contains a circle with radius R_m and is contained in a circle with radius R_M; if ϱ_M resp. ϱ_m denote the radii of the least circumscribable resp. the greatest inscribable circle of $C(r)$ it follows $R_m\leqq\varrho_m<\varrho_M\leqq R_M$, and thus, using (32) we obtain

Corollary II.

$$\varrho_M-\varrho_m=O(r^3). \tag{33}$$

As remarked in the introduction, the distortion theorem gives only $\varrho_M-\varrho_m=O(r^2)$. (Of course the least circumscribable and the greatest inscribable circle are generally not concentric.)

Finally we mention that R_M always exceeds r. This follows from the fact, that $\dfrac{R(r,\varphi)}{r}$ is a subharmonic function. As a matter of fact, it suffices to show that $\log\dfrac{R}{r}$ is subharmonic. As regards the latter function, we have

$$\log\frac{R}{r}=R(\log f'(z))-\log R\left(1+\frac{zf''(z)}{f'(z)}\right). \tag{34}$$

The first term on the right of (34) is a harmonic function, and the second — being the negative logarithm of a harmonic function — is subharmonic, and thus $\log\dfrac{R}{r}$, and therefore also $\dfrac{R}{r}$ itself are subharmonic; as the maximal value of a subharmonic function can not be taken in an interior point and as $\dfrac{R}{r}=1$ for $r=0$, it follows $R_M>r$.

Part II.

Let $r(\lambda)$ $(0\leqq\lambda\leqq 1)$ denote the least upper bound of those values of r for which, for any $f(z)\in S$, the star-kernel $K(r)$ contains $D(\lambda r)$. According to the theorems on the radii of convexity and starlikeness, cited in the introduction, we have $r(1)=2-\sqrt{3}$ and $r(0)=\tanh\dfrac{\pi}{4}$. Evidently $r(\lambda)$ is a continuous decreasing function of λ. In what follows we shall prove the following estimate for $r(\lambda)$:

[10]) W. BLASCHKE, *Kreis und Kugel* (Leipzig, 1916), p. 115.

Theorem 2. *We have for* $0<\lambda<\frac{\pi-\log 3}{2e^{\pi/2}}$

$$r(\lambda)>\tanh\left(\frac{\pi}{4}-\frac{e^{\pi/2}}{2}\lambda\right). \tag{35}$$

Proof. It is easy to see that

$$\left|\arg\frac{zf'(z)}{f(z)-f(a)}\right|\leq\frac{\pi}{2} \tag{36}$$

for $z=re^{i\varphi}$, $0\leq\varphi<2\pi$, is the necessary and sufficient condition for $C(r)$ being star-like with respect to the point $w=f(a)$. Let us put $\zeta=\frac{a-z}{1-a\bar z}$; it follows $a=\frac{\zeta+z}{1+\bar z\zeta}$.

We need the following theorem, valid for any $f(z)\in S$, which has been proved first by GRUNSKY:[11])

$$\left|\arg\frac{f(z)}{z}\right|\leq\log\frac{1+|z|}{1-|z|}. \tag{37}$$

Let us apply (37) to the function $h(\zeta)$ defined by (16), we obtain

$$\left|\arg\frac{f(a)-f(z)}{f'(z)\zeta}\right|\leq\log\frac{1+|\zeta|}{1-|\zeta|} \tag{38}$$

and thus

$$\left|\arg\frac{zf'(z)}{f(z)-f(a)}\right|=\left|\arg\left(\frac{f(a)-f(z)}{f'(z)\zeta}\right)\cdot\left(-\frac{\zeta}{z}\right)\right|\leq\log\frac{1+|\zeta|}{1-|\zeta|}+\left|\arg\left(-\frac{\zeta}{z}\right)\right|. \tag{39}$$

The circle $|z|=r$ is mapped by $\zeta_1=\frac{z-a}{1-\bar a z}$ onto the circle with centre $-a\frac{1-r^2}{1-r^2|a|^2}$ and radius $\frac{r(1-|a|^2)}{1-r^2|a|^2}$. As $|\zeta_1|=|\zeta|$ it follows that for $|z|=r$ and for any a with $|a|=\varrho$ we have $|\zeta|\leq\frac{\varrho+r}{1+\varrho r}$. We have further for $|z|=r$ and $|a|=\varrho$

$$\left|\arg\left(-\frac{\zeta}{z}\right)\right|=\left|\arg\frac{1-\frac{a}{z}}{1-a\bar z}\right|\leq\operatorname{arctg}\frac{\varrho}{r}+\operatorname{arctg}\varrho r=\operatorname{arctg}\frac{\varrho\left(r+\frac{1}{r}\right)}{1-\varrho^2}.$$

Thus it follows that for $|z|=r$ and $|a|\leq\lambda r$ $(\lambda>0)$

$$\left|\arg\frac{zf'(z)}{f(z)-f(a)}\right|\leq\log\frac{1+r}{1-r}+\log\frac{1+\lambda r}{1-\lambda r}+\operatorname{arctg}\frac{\lambda(r^2+1)}{1-\lambda^2 r^2}. \tag{40}$$

Using the elementary inequalities $\log\frac{1+x}{1-x}\leq\frac{2x}{1-x^2}$ and $\operatorname{arctg}x\leq x$ we obtain

[11]) See H. GRUNSKY, l. c. [2]).

$$\left|\arg \frac{zf'(z)}{f(z)-f(a)}\right| \leq \log \frac{1+r}{1-r} + \frac{\lambda(r+1)^2}{1-\lambda^2 r^2} \leq \log \frac{1+r}{1-r} + \lambda \frac{1+r}{1-r}. \tag{41}$$

Taking into account that we may suppose $r < \tanh \frac{\pi}{4}$, we obtain

$$\left|\arg \frac{zf'(z)}{f(z)-f(a)}\right| \leq \log \frac{1+r}{1-r} + \lambda e^{\frac{\pi}{2}} \tag{42}$$

thus if $r \leq \tanh\left(\frac{\pi}{4} - \frac{\lambda e^{\frac{\pi}{2}}}{2}\right)$, i. e. if $\log \frac{1+r}{1-r} \leq \frac{\pi}{2} - \lambda e^{\frac{\pi}{2}}$, we have for any a with $|a| \leq \lambda r$

$$\left|\arg \frac{zf'(z)}{f(z)-f(a)}\right| \leq \frac{\pi}{2} \tag{43}$$

which proves Theorem 2.

We may deduce from (41) also the slightly more precise result

$$r(\lambda) \geq \frac{x(\lambda)-1}{x(\lambda)+1} \tag{44}$$

where $x(\lambda)$ is the only positive root of the equation

$$\log x + \lambda x = \frac{\pi}{2}. \tag{45}$$

These estimations are not the best possible, nevertheless they give rather good approximation for small values of λ.

(Received August 6, 1949, revised November 10, 1949.)

Sur la convergence et la sommabilité presque partout des séries de polynomes orthogonaux.

Par GEORGES ALEXITS à Budapest.

1. Étant donné dans l'intervalle fermé $[-1, 1]$ une fonction non-négative $w(x)$ intégrable, elle y détermine — comme on sait — exactement un système $\{p_n(x)\}$ de polynomes orthogonaux et normés. Soit

$$(1)\qquad f(x) \sim \sum_{n=0}^{\infty} c_n p_n(x)$$

où

$$c_n = \int_{-1}^{+1} w(x) f(x) p_n(x)\, dx,$$

le développement formel suivant les polynomes $p_n(x)$ d'une fonction $f(x)$.

On ne sait que très peu sur la convergence et la sommabilité presque partout du développement (1), excepté le cas où la nature de la fonction de poids $w(x)$ est restreinte par des conditions de continuïté assez étroites qui permettent de ramener ces problèmes aux problèmes analogues concernant le développement en série de Fourier de la fonction $f(\cos\theta)$ où $\theta = \arccos x$. Dans ce qui suit, nous allons rechercher ces problèmes sans supposer beaucoup sur la nature de la fonction $w(x)$, mais en introduisant certaines conditions concernant les coefficients de Fourier de la fonction $f(\cos\theta)$. Il en résulte que, même si l'on ne sait rien sur l'équiconvergence de la série (1) et de la série de Fourier de $f(\cos\theta)$, il y a une analogie entre le comportement de ces deux espèces de développements et on peut dire que la série de Fourier de $f(\cos\theta)$ joue, en certain sens, un rôle déterminant dans le problème de la convergence et de la sommabilité presque partout du développement (1).

2. En introduisant, au lieu de x, la variable $\theta = \arccos x$, la fonction $f(\cos\theta)$ définie dans l'intervalle $[-\pi, \pi]$ est paire, la n-ième somme partielle

[1]) En ce qui concerne les théorèmes respectifs, voir par exemple G. SZEGŐ, *Orthogonal Polynomials* (New York, 1939).

de son développement de Fourier est donc de la forme

$$s_n(\theta)=\frac{a_0}{2}+\sum_{k=1}^{n} a_k \cos k\theta$$

où

$$a_k=\frac{2}{\pi}\int_0^{\pi} f(\cos\theta)\cos k\theta\, d\theta \qquad (k=0, 1, \ldots).$$

Lemme.[2]) *Si* $0\leqq w(x)\leqq W(1-x^2)^{-\frac{1}{2}}$ *où* W *est une constante arbitraire, on a*

$$\sum_{k=n}^{\infty} c_k^2\leq\frac{W\pi}{2}\sum_{k=n}^{\infty} a_k^2 \tag{2}$$

et, d'une manière plus générale,

$$\sum_{k=1}^{\infty} \lambda_k c_k^2\leq\frac{W\pi}{2}\sum_{k=1}^{\infty} \lambda_k a_k^2 \tag{3}$$

pour toute suite croissante $\{\lambda_n\}$ *de nombres positifs.*

Démonstration. Désignons par $S_n(x)$ la n-ième somme partielle du développement (1). Il est connu que

$$\int_{-1}^{+1} w(x)\left[f(x)-S_{n-1}(x)\right]^2 dx\leqq\int_{-1}^{+1} w(x)\left[f(x)-P_{n-1}(x)\right]^2 dx,$$

quel que soit le polynome $P_{n-1}(x)$ de degré $n-1$. Choisissons pour $P_{n-1}(x)$ le polynome de degré $n-1$ qu'on obtient de $s_{n-1}(\theta)$ en y posant $\theta=\arc\cos x$. Il s'ensuit de l'inégalité précédente:

$$\sum_{k=n}^{\infty} c_k^2=\int_{-1}^{+1} w(x)\left[f(x)-S_{n-1}(x)\right]^2 dx\leq\int_0^{\pi} w(\cos\theta)\sin\theta\left[f(\cos\theta)-s_{n-1}(\theta)\right]^2 d\theta\leq$$

$$\leq W\int_0^{\pi}\left[f(\cos\theta)-s_{n-1}(\theta)\right]^2 d\theta=\frac{W\pi}{2}\sum_{k=n}^{\infty} a_k^2$$

et l'inégalité (2) est démontrée.

Pour démontrer l'inégalité (3), désignons les deux membres de (2) par C_n et A_n. On a $c_k^2=C_k-C_{k+1}$, $\frac{W\pi}{2}a_k^2=A_k-A_{k+1}$. D'après (2), on a $C_n\leq A_n$;

[2]) Qu'il me soit permis d'exprimer mes remerciements à M. Béla Sz.-Nagy à qui je dois cette forme de mon lemme et un raccourcissement de ma démonstration originelle. Sa remarque permet aussi une amélioration de mes résultats antérieurs concernant la convergence absolue et l'ordre d'approximation des développements en série de polynomes orthogonaux (G. Alexits, Sur la convergence des séries de polynomes orthogonaux, *Commentarii Math. Helvetici*, **16** (1943), p. 200—208). On peut notamment montrer que mes résultats respectifs restent exactes même si l'on y remplace l'hypothèse originelle $0\leq w(x)\leqq W$ par l'hypothèse plus large $0\leqq w(x)\leqq W(1-x^2)^{-\frac{1}{2}}$.

il en résulte par une transformation abélienne (en posant $\lambda_0=0$):

$$\sum_{k=1}^{n-1}\lambda_k c_k^2=\sum_{k=1}^{n}(\lambda_k-\lambda_{k-1})C_k-\lambda_n C_n\leqq\sum_{k=1}^{n}(\lambda_k-\lambda_{k-1})A_k=$$

$$=\frac{W\pi}{2}\sum_{k=1}^{n-1}\lambda_k a_k^2+\lambda_n A_n\leqq\frac{W\pi}{2}\left(\sum_{k=1}^{n-1}\lambda_k a_k^2+\sum_{k=n}^{\infty}\lambda_k a_k^2\right),$$

ce qui prouve (3).

3. Théorème 1. *Si* $0\leqq w(x)\leqq W(1-x^2)^{-1/2}$ *et si* $\sum a_n^2\log^2 n<\infty$ *respectivement* $\sum a_n^2(\log\log n)^2<\infty$, *la série* (1) *est presque partout convergente, respectivement presque partout sommable* $(C,1)$.

Pour le démontrer, on n'a qu'à appliquer le lemme avec $\lambda_n=\log^2 n$, respectivement avec $\lambda_n=(\log\log n)^2$ et à rappeler que la convergence des séries

$$\sum c_n^2\log^2 n,\quad \text{resp.}\quad \sum c_n^2(\log\log n)^2$$

entraîne la convergence, respectivement la sommabilité $(C,1)$ presque partout de la série (1)[3].

Théorème 2. *Si* $0\leqq w(x)\leqq W(1-x^2)^{-1/2}$ *est si les polynomes* $p_n(x)$ *sont uniformément bornés, la convergence de la série* $\sum a_n^2\log n$ *entraîne la convergence presque partout de la série* (1).

En posant $\lambda_n=\log n$, on obtient de (3) la convergence de la série $\sum c_n^2\log n$. Or ce fait, combiné avec l'hypothèse que les $p_n(x)$ soient uniformément bornés, implique, d'après un résultat antérieur de l'auteur[4], la convergence presque partout de la série (1).

(Reçu le 31 décembre 1949)

[3]) S. Kaczmarz—H. Steinhaus, *Theorie der Orthogonalreihen* (Warszawa—Lwów, 1935), p. 164, [535] et p. 190, [586].

[4]) L. c.[2]), p. 201.

On the mapping of the unit-circle by polynomials.

By E. EGERVÁRY in Budapest.

1. Consider the set $\{\Pi\}$ of rational polynomials $P(z)$, which have the following properties:

1. The degree of $P(z)$ is not higher than n,
2. $P(0)=0$,
3. $\Re P(z) \geqq -1$ for $|z| \leqq 1$.

As an application of his theory of non-negative trigonometrical polynomials L. FEJÉR has proved that[1])

$$-1 \leqq \Re P(e^{it}) \leqq n. \tag{1}$$

Somewhat later O. SZÁSZ proved the complementary inequality[2])

$$-\operatorname{ctg}\frac{\pi}{2(n+1)} \leqq \Im P(e^{it}) \leqq \operatorname{ctg}\frac{\pi}{2(n+1)}. \tag{2}$$

Hence all the maps of the unit-circle generated by the polynomials $P(z)$ of the set $\{\Pi\}$ lie in the rectangle

$$-1 \leqq \Re w \leqq n, \quad -\operatorname{ctg}\frac{\pi}{2(n+1)} \leqq \Im w \leqq \operatorname{ctg}\frac{\pi}{2(n+1)}.$$

In the present article I wish to determine the precise field of variability of the maps of the unit-circle which are generated by the set $\{\Pi\}$. The chief results may be stated in the following theorems:

Theorem I. *The point-theoric sum Σ of the maps of the unit-circle which are generated by the set $\{\Pi\}$ is a convex region which coincides with the convex hull of the map of the unit-circle generated by the polynomial*

$$P^*(z)=\frac{2}{n+1}\{nz+(n-1)z^2+\ldots+1\,.\,z^n\}, \quad P^*(z)\subset\{\Pi\}. \tag{3}$$

Theorem II. *The supporting function of Σ with respect to* 0 *is given by*

$$p(\theta)=\sin\frac{n\theta}{n+1}\Big/\sin\frac{\theta}{n+1}, \quad -\pi\leqq\theta\leqq\pi. \tag{4}$$

---

1) L. FEJÉR, Über trigonometrische Polynome, *Journal für die reine und angewandte Math.*, **146** (1913), pp. 53–82.

2) O. SZÁSZ, Über harmonische Funktionen und L-Formen, *Math. Zeitschrift,* **1** (1918), pp. 149–162.

For the special values $\theta=0$ and $\theta=\frac{\pi}{2}$ this yields the above-mentioned results of FEJÉR and SZÁSZ.

The excentricity of a bounded region with respect to an interior point C may be measured by the ratio of the maximum and minimum of the supporting function of its convex hull with respect to C. Adopting this definition we deduce from our former results the

Theorem III. *The excentricity of the map of the unit-circle generated by an arbitrary polynomial $Q(z)$ of degree n with respect to $Q(0)$ cannot exceed n, this maximal excentricity being attained only for $a+bP^*(\varepsilon z)$ (a, b are arbitrary complex numbers, $|\varepsilon|=1$).*

2. The proof of these theorems will be based on the following theorem of L. FEJÉR[1]).

The set $\{\Pi\}$ of the polynomials $P(z)$ of degree n

$$P(z)=c_1z+c_2z^2+\ldots+c_nz^n$$

which map the unit-circle on the half-plane $\Re(z)\geqq-1$, admits the representation

$$P(z)=\sum_{k=1}^{n}(\bar{\gamma}_0\gamma_k+\bar{\gamma}_1\gamma_{k+1}+\ldots+\bar{\gamma}_{n-k}\gamma_n)z^k, \tag{5}$$

where $\gamma_0, \gamma_1 \ldots \gamma_n$ are arbitrary complex parameters subjected only to the restriction

$$|\gamma_0|^2+|\gamma_1|^2+\ldots+|\gamma_n|^2=1. \tag{5'}$$

Hence if $a_1, a_2, \ldots, a_n$ denote arbitrary constants, the maximum (minimum) of the linear form

$$\Re\sum_{k=1}^{n}a_kc_k, \quad \sum_{k=1}^{n}c_kz^k\subset\{\Pi\}$$

is equal to the greatest (least) root λ^* of the characteristic equation

$$\begin{vmatrix} -\lambda & a_1 & a_2 & \ldots & a_n \\ \bar{a}_1 & -\lambda & a_1 & \ldots & a_{n-1} \\ \bar{a}_2 & \bar{a}_1 & -\lambda & \ldots & a_{n-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \bar{a}_n & \bar{a}_{n-1} & \bar{a}_{n-2} & \ldots & -\lambda \end{vmatrix}=0.$$

The maximum (minimum) of $\Re\sum_{1}^{n}a_kc_k$ will be attained for the polynomial $P^*(z)=\sum_{1}^{n}(\bar{\gamma}_0^*\gamma_k^*+\bar{\gamma}_1^*\gamma_{k+1}^*+\ldots+\bar{\gamma}_{n-k}^*\gamma_n^*)z^k$, where $\gamma_0^*, \gamma_1^*, \ldots, \gamma_n^*$ are determined by the system of equations

$$(6)\qquad \begin{aligned} -\lambda^*\gamma_0^* + a_1\gamma_1^* + \ldots + a_n\gamma_n^* &= 0\\ \bar{a}_1\gamma_0^* - \lambda^*\gamma_1^* + \ldots + a_{n-1}\gamma_n^* &= 0\\ \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot&\\ \bar{a}_n\gamma_0^* + \bar{a}_{n-1}\gamma_1^* + \ldots - \lambda^*\gamma_n^* &= 0\\ |\gamma_0^*|^2 + |\gamma_1^*|^2 + \ldots + |\gamma_n^*|^2 &= 1. \end{aligned}$$

3. We prove first the following

Lemma. *The point-theoric sum Σ of all the maps of the unit-circle generated by the polynomials $P(z)$ of the set $\{\Pi\}$ is a convex region.*

Indeed, if w_1 and w_2 belong to the region Σ then there are polynomials $P_1(z)$ and $P_2(z)$ such that

$$P_1(z) \subset \{\Pi\},\ w_1 = P_1(z_1),\ |z_1| \leqq 1,$$
$$P_2(z) \subset \{\Pi\},\ w_2 = P_2(z_2),\ |z_2| \leqq 1.$$

Consider now the polynomial $Q(z) = \dfrac{P_1(z_1 z) + \mu P_2(z_2 z)}{1+\mu}$, $\mu > 0$. Obviously $Q(z) \subset \{\Pi\}$ and $Q(1) = \dfrac{P_1(z_1) + \mu P_2(z_2)}{1+\mu} = \dfrac{w_1 + \mu w_2}{1+\mu}$, consequently, if w_1 and w_2 belong to the region Σ, then the segment of straight line joining them is contained in Σ, q. e. d.

4. In order to determine the supporting function $p(\theta)$ of the convex region Σ with respect to 0, I will make use of the following representation of the supporting function

$$(7)\qquad p(\theta) = \max_{w \subset \Sigma} \Re\{e^{-i\theta} w\}, \qquad -\pi \leqq \theta \leqq \pi.$$

The set of the points $w \subset \Sigma$ is given by

$$w = P(z);\ P(z) \subset \{\Pi\},\ |z| \leqq 1,$$

hence

$$p(\theta) = \max \Re\{e^{-i\theta} P(z)\},\ P(z) \subset \{\Pi\},\ |z| \leqq 1.$$

But the harmonic function $\Re\{e^{-i\theta}P(z)\}$ attains its extremal values on the boundary, therefore

$$p(\theta) = \max \Re\{e^{-i\theta} P(e^{it})\},\ P(z) \subset \{\Pi\},\ -\pi \leqq t \leqq \pi.$$

Suppose, that the maximum will be attained for a $P^*(z) \subset \{\Pi\}$ and for $z = e^{it_1}$. Then $P^{**}(z) = P^*(ze^{it_1}) \subset \{\Pi\}$ and $P^{**}(1) = P^*(e^{it_1})$, consequently

$$p(\theta) = \max \Re\{e^{-i\theta} P(1)\},\ P(z) \subset \{\Pi\}.$$

The set of the polynomials $P(z)$ belonging to $\{\Pi\}$ is given by

$$P(z) = \sum_1^n (\bar{\gamma}_0\gamma_k + \bar{\gamma}_1\gamma_{k+1} + \ldots + \bar{\gamma}_{n-k}\gamma_n) z^k,\quad \sum_0^n |\gamma_h|^2 = 1$$

hence

$$p(\theta) = \max \Re\{e^{-i\theta} P(1)\} = \max \Re\left\{e^{i\theta} \sum_1^n (\bar{\gamma}_0\gamma_k + \ldots + \bar{\gamma}_{n-k}\gamma_n)\right\};\ \sum_0^n |\gamma_k|^2 = 1.$$

We conclude herefrom that the supporting function $p(\theta)$ will be given by the greatest root of the characteristic equation

$$\begin{vmatrix} -\lambda & e^{i\theta} & e^{i\theta} & \dots & e^{i\theta} \\ e^{-i\theta} & -\lambda & e^{i\theta} & \dots & e^{i\theta} \\ e^{-i\theta} & e^{-i\theta} & -\lambda & \dots & e^{i\theta} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ e^{-i\theta} & e^{-i\theta} & e^{-i\theta} & \dots & -\lambda \end{vmatrix} = (-1)^{n+1} \frac{e^{-i\theta}(\lambda+e^{i\theta})^{n+1} - e^{i\theta}(\lambda+e^{-i\theta})^{n+1}}{e^{-i\theta}-e^{i\theta}} = 0.^{3)}$$

The roots of this equation are

$$\lambda_k(\theta) = \sin\frac{n\theta - k\pi}{n+1} \Big/ \sin\frac{\theta+k\pi}{n+1} \qquad (k=0, 1, \dots, n),$$

and it is obvious that

$$\lambda_0(\theta) > \lambda_1(\theta) > \dots > \lambda_n(\theta) \quad \text{for} \quad -\pi \leqq \theta \leqq \pi.$$

Consequently, the supporting function of the region Σ is given by

$$(8) \qquad p(\theta) = \lambda_0(\theta) = \sin\frac{n\theta}{n+1} \Big/ \sin\frac{\theta}{n+1}, \qquad -\pi \leqq \theta \leqq \pi,$$

and an easy calculation shows that the boundary of the convex region Σ consists of the arc $-\frac{2\pi}{n+1} \leqq t \leqq \frac{2\pi}{n+1}$ of the curve

$$(9) \qquad w = \frac{2}{n+1} \frac{n e^{it} - (n+1) e^{2it} + e^{(n+2)it}}{(1-e^{it})^2}$$

and of the segment $-\operatorname{ctg}\frac{\pi}{n+1} \leqq t \leqq \operatorname{ctg}\frac{\pi}{n+1}$ of the straight line

$$(10) \qquad w = -1 + it.$$

By the determination of the extremal polynomials we avoid the direct solution of the equations (6) and proceed as follows.

Consider the arithmetic mean

$$P^*(z) = \frac{2}{n+1} \{nz + (n-1) z^2 + \dots + 1 \,.\, z^n\}$$

of the partial sums of the geometrical series

$$\frac{2z}{1-z} = 0 + 2z + 2z^2 + 2z^3 + \dots$$

This polynomial maps the unit-circle $|z| \leqq 1$ on a simple, starshaped region, whose supporting function (with respect to 0) is given by

$$\max_{|z| \leqq 1} \Re\{e^{-i\theta} P^*(z)\} = \Re\left\{e^{-i\theta} P^*\left(e^{\frac{2i\theta}{n+1}}\right)\right\} = \sin\frac{n\theta}{n+1} \Big/ \sin\frac{\theta}{n+1} = p(\theta).$$

[3]) See f. i. PÓLYA—SZEGŐ, *Aufgaben und Lehrsätze aus der Analysis*, II (Berlin, 1925), p. 99.

Consequently, the map generated by $w = P^*(z)$ of the arc $z = e^{it}$ $-\frac{2\pi}{n+1} \leqq t \leqq \frac{2\pi}{n+1}$ coincides with the arc (9) of the boundary of the convex region Σ.

We have proved in this way that the region Σ is identical to the convex hull of the curve $w = P^*(e^{it})$ on which the unit-circle $z = e^{it}$ is mapped by the polynomial

$$P^*(z) = \frac{2}{n+1}\{nz + (n-1)z^2 + \ldots + 1 \,.\, z^n\}.$$

Those points of the boundary of Σ which belong to the line-segment (10) will be attained for

$$Q^*(z) = \frac{P^*\left(ze^{\frac{2\pi i}{n+1}}\right) + \mu P^*\left(ze^{-\frac{2\pi i}{n+1}}\right)}{1+\mu} \subset \{\Pi\}; \quad z = 1, 0 < \mu < \infty.$$

5. It is evident that the supporting function of a convex region with respect to a point C cannot be constant unless the region is a circle whose center coincides with C. Hence it seems to be natural to define the excentricity of a region (with respect to an interior point) as the ratio of the maximum and minimum of the supporting function belonging to its convex hull.

Now let $Q(z)$ be an *arbitrary* polynomial of degree n. Replace $Q(z)$ by $P(z) = \Gamma\{Q(\varepsilon z) - Q(0)\}$ where Γ and ε are so chosen that

$$\min_{|\theta| \leqq \pi} [\max_{|z| \leqq 1} \Re\{e^{i\theta} P(z)\}] = \Re\{P(1)\} = -1$$

consequently

$$\Re\{P(z)\} \geqq -1 \quad \text{for} \quad |z| \leqq 1.$$

From our former results we infer that the supporting function $p(\theta)$ belonging to the convex hull of the region $w = \Gamma\{Q(\varepsilon z) - Q(0)\}$, $|z| \leqq 1$ verifies the following conditions: $p(0) = \min p(\theta) = 1$ and

$$\max p(\theta) \leqq \max \left\{ \sin\frac{n\theta}{n+1} \Big/ \sin\frac{\theta}{n+1} \right\} = n = n \, \min p(\theta).$$

But the excentricity of a region is obviously invariant under the homothetic transformation $P(z) = \Gamma\{Q(\varepsilon z) - Q(0)\}$. Thus the inequality

$$\max p(\theta) \leqq n \, \min p(\theta)$$

is generally proved.

(Received January 17, 1950.)

Sur quelques propriétés des dérivées des fonctions d'une variable réelle.

Par NIKOLA OBRECHKOFF à Sofia.

Dans ce travail nous démontrons quelques inégalités pour les dérivées des fonctions réelles définies sur le demi-axe ou sur tout l'axe réel et des inégalités pour les différences des suites de nombres réels. Nous en déduirons aussi quelques propriétés nouvelles pour les fonctions réelles.

1. Fonctions définies sur le demi-axe réel.

Théorème I. *Soit $f(x)$ une fonction réelle telle que $f^{(n)}(x) \geqq 0$ pour $x > a$. Supposons qu'il existe une suite*

$$(1) \qquad \{y_\lambda\}_1^\infty \qquad \lim_{\lambda \to \infty} y_\lambda = \infty,$$

et un entier m $(0 \leqq m < n)$ tels que

$$(2) \qquad \lim_{\lambda \to \infty} \frac{f(y_\lambda)}{y_\lambda^m} = 0.$$

On a alors pour $x > a$

$$(3) \qquad (-1)^{n-m} f^{(m)}(x) \geqq 0, \ (-1)^{n-m-1} f^{(m+1)}(x) \geqq 0, \ldots, \ -f^{(n-1)}(x) \geqq 0,$$

$$(4) \qquad \lim_{x \to \infty} \frac{f^{(i)}(x)}{x^{m-i}} = 0 \quad (0 \leqq i \leqq m-1), \quad \lim_{x \to \infty} f^{(i)}(x) = 0 \quad (m \leqq i \leqq n-1).$$

De plus, si la fonction $f^{(i)}(x)$ $(m \leqq i \leqq n-1)$ s'annule pour un $x = b > a$, elle s'annule pour tout $x > b$.

Supposons que le théorème soit déjà démontré pour $n-1$. Puisque $f^{(n)}(x) \geqq 0$, la fonction $f^{(n-1)}(x)$ est non décroissante pour $x > a$. Supposons que pour un nombre α on ait $f^{(n-1)}(\alpha) = C > 0$; alors $f^{(n-1)}(x) \geqq C$ pour $x \geqq \alpha$. On en obtient par intégration que

$$f(x) \geqq C \frac{x^{n-1}}{(n-1)!} + P(x)$$

où $P(x)$ est un polynome de degré $\leqq n-2$. Donc on aura

$$\lim_{x \to \infty} \frac{f(x)}{x^{n-1}} \geqq \frac{C}{(n-1)!} > 0,$$

ce qui est en contradiction avec (2). Par conséquent $f^{(n-1)}(x) \leqq 0$ pour $x > a$. Si $m = n-1$, la première des inégalités (3) est démontrée. Soit $m < n-1$; le théorème étant vrai pour $n-1$, on a

$$(-1)^{n-1-m}[-f(x)]^{(m)} \geqq 0, \text{ donc } (-1)^{n-m} f^{(m)}(x) \geqq 0 \text{ pour } x > a.$$

Les autres inégalités de (3) découlent d'ici puisque, par (2),

$$\lim_{\lambda \to \infty} \frac{f(y_\lambda)}{y_\lambda^k} = 0 \qquad \text{pour} \quad k \geqq m+1.$$

Si $n-m$ est pair, les deux premières des inégalités (3) assurent que la fonction $f^{(m)}(x)$ soit non négative et non croissante. Par conséquent, elle tend vers une limite B lorsque $x \to \infty$. D'après la règle d'Hospital, on aura $\lim\limits_{x \to \infty} (f(x)/x^m) = m!\,B$, d'où et de (2) il suit que $B = 0$. Si $n-m$ est impair, $f^{(m)}(x)$ est non positive et non décroissante et on a le même résultat. Les autres égalités de (4) découlent d'ici immédiatement.

Supposons enfin que $f^{(i)}(b) = 0$ pour un i, $m \leqq i < n$. La fonction $(-1)^{n-i} f^{(i)}(x)$ étant, en vertu des inégalités (3), non négative et non croissante pour $x > a$, s'annule nécessairement pour $x \geqq b$. Cela achève la démonstration du théorème.

Soient maintenant $\varphi(x)$ et $\psi(x)$ deux fonctions réelles qui admettent pour $x > a$ les dérivées $\varphi^{(n)}(x)$ et $\psi^{(n)}(x)$ et supposons que pour une suite (1) et pour un nombre entier m $(0 \leqq m < n)$ les limites

$$\lim_{\lambda \to \infty} \frac{\varphi(y_\lambda)}{y_\lambda^m} = B, \quad \lim_{\lambda \to \infty} \frac{\psi(y_\lambda)}{y_\lambda^m} = C \tag{7}$$

existent. Si

$$\psi^{(n)}(x) \leqq \varphi^{(n)}(x) \qquad (x > a), \tag{6}$$

on aura

$$(-1)^{n-m}[\psi^{(m)}(x) - m!\,C] \leqq (-1)^{n-m}[\varphi^{(m)}(x) - m!\,B] \qquad (x > a). \tag{7}$$

Cette proposition découle immédiatement du théorème I, en considérant la fonction auxilière $f(x) = \varphi(x) - \psi(x) - (B-C)\,x^m$ pour laquelle $\lim f(y_\lambda)/y_\lambda^m = 0$ et $f^{(n)}(x) \geqq 0$. La remarque pour le signe d'égalité dans (7) reste valable.

Si au lieu de (6) on a

$$|\psi^{(n)}(x)| \leqq |\varphi^{(n)}(x)|, \tag{8}$$

la fonction $\varphi^{(n)}(x)$ ne changeant pas de signe pour $x > a$, on aura au lieu de (7) l'inégalité suivante

$$|\psi^{(m)}(x) - m!\,C| \leqq |\varphi^{(m)}(x) - m!\,B|.$$

En effet, si par exemple $\varphi^{(n)}(x) \geqq 0$ pour $x > a$, l'inégalité (8) est équivalente à $-\varphi^{(n)}(x) \leqq \psi^{(n)}(x) \leqq \varphi^{(n)}(x)$ et on aura

$$(-1)^{n-m-1}[\varphi^{(m)}(x) - m!\,B] \leqq (-1)^{n-m}[\psi^{(m)}(x) - m!\,C] \leqq (-1)^{n-m}[\varphi^{(m)}(x) - m!\,B].$$

On a la même remarque pour le signe d'égalité.

2. Fonctions définies sur tout l'axe réel.

Théorème II. *Soit $f(x)$ une fonction réelle telle que $f^{(n)}(x) \geqq 0$ pour $-\infty < x < \infty$. Supposons encore qu'il existe une suite à deux côtés*

(9) $$\{y_\lambda\}_{-\infty}^{+\infty}, \quad \lim_{\lambda\to-\infty} y_\lambda = -\infty, \quad \lim_{\lambda\to\infty} y_\lambda = \infty,$$

et un entier m $(0 \leq m < n)$ tels que

(10) $$\lim_{\lambda\to\pm\infty} \frac{f(y_\lambda)}{y_\lambda^m} = 0.$$

Alors la fonction $f(x)$ est un polynome de degré $\leq m-1$.

En effet, de la condition (10) pour $\lambda \to \infty$ et du théorème I il suit que $f^{(n-1)}(x) \leq 0$ pour tous les x. Considérons maintenant la fonction $\varphi(x) = (-1)^n f(-x)$. On a $\varphi^{(n)}(x) = f^{(n)}(-x) \geqq 0$ pour $-\infty < x < \infty$. De la condition (10) pour $\lambda \to -\infty$ et du théorème I on conclut alors que $f^{(n-1)}(-x) \geqq 0$. Donc $f^{(n-1)}(x) = 0$ pour tous les x et $f(x)$ est un polynome dont le degré, à cause de (10), ne surpasse pas le nombre $m-1$.

Ce théorème a les corollaires suivants:

a) *Si $f^{(n)}(x) \geqq 0$ et $|f(x)| < K(1+|x|^m)$ pour $-\infty < x < \infty$, alors $f(x)$ est un polynome de degré $\leqq m$.*

b) *Soit $\varphi(x) \geqq 0$ pour $-\infty < x < \infty$ et soit $f(x)$ une autre fonction réelle qui pour $-\infty < x < \infty$ admet des dérivées jusqu'à l'ordre n. Supposons encore que pour chaque x réel on ait $\frac{d^n}{dx^n}\frac{f(x)}{\varphi(x)} \geqq 0$ et $|f(x)| < K\varphi(x)$, K étant une constante. Alors $f(x) = K_1\varphi(x)$ où K_1 est une constante.*

Prenant en particulier $\varphi(x) = e^x$, on obtient la proposition suivante:

Les inégalités $|f(x)| < Ke^x$ et $(-1)^n \sum_{\nu=1}^{n} \binom{n}{\nu} (-1)^\nu f^{(\nu)}(x) \geqq 0$ pour chaque x entraînent que $f(x) = K_1 e^x$, K_1 étant une constante.

Théorème III. *Soit $f(x)$ une fonction réelle telle que $f^{(2n)}(x) \geqq 0$ pour $-\infty < x < \infty$. Supposons encore que pour une suite infinie $\{y_\lambda\}_{-\infty}^{\infty}$ de type (9) on ait $\lim_{\lambda\to\pm\infty} f(y_\lambda)/y_\lambda^{2n-1} = 0$. Alors $f(x)$ est un polynome de degré $2n-2$.*

La fonction $f^{(2n-1)}(x)$ est non décroissante. Supposons que pour un α on ait $f^{(2n-1)}(\alpha) = C > 0$. Alors on conclut comme plus haut que pour $x > \alpha$ on a $f(x) > \frac{C}{(2n-1)!}x^{2n-1} + R(x)$ où $R(x)$ est un polynome de degré $< 2n-1$, ce qui est en contradiction aves nos hypothèses. Donc on aura $f^{(2n-1)}(x) \leq 0$ pour tous les x. Considérons maintenant la fonction $\varphi(x) = f(-x)$. Comme on a $\varphi^{(2n)}(x) = f^{(2n)}(-x) \geqq 0$, le même raisonnement vérifie que $\varphi^{(2n-1)}(x) \leq 0$, c'est-à-dire $f^{(2n-1)}(x) \geqq 0$ pour tous les x. Donc $f^{(2n-1)}(x) = 0$ et le théorème est démontré.

Théorème IV. *Soit $\psi(x)$ une fonction réelle qui admet pour tous les x les dérivées jusqu'à l'ordre $2n$ et telle que*

$$\sum_{\nu=0}^{2n}(-1)^{\nu}\binom{2n}{\nu}\psi^{(\nu)}(x)\geqq 0 \tag{11}$$

pour tous les x. Supposons encore que pour une suite $\{y_\lambda\}_{-\infty}^{+\infty}$ de type (9) *on ait*

$$\psi(y_\lambda)<Q(y_\lambda)\,e^{y_\lambda}$$

où $Q(x)$ est un polynome de degré $2n-2$. La fonction $\psi(x)$ est alors égale à $Q_1(x)e^x$ où $Q_1(x)$ est un polynome de degré $2n-2$.

Ce théorème découle immédiatement du précédent en l'appliquant à la fonction $f(x)=\psi(x)\,e^{-x}$.

D'après S. Bernstein, une fonction réelle $f(x)$ est dite absolument monotone dans un intervalle (a,b) si elle y est indéfiniment dérivable et si

$$f^{(n)}(x)\geqq 0 \qquad (a<x<b;\ n=0,1,2,\dots).$$

On sait bien que pour une telle fonction on a $f''(x)-2f'(x)+f(x)\geqq 0$. Donc on obtient du théorème II, comme cas particulier, le résultat suivant:

Si la fonction $f(x)$ est absolument monotone dans $(-\infty,\infty)$ et si pour une suite $\{y_\lambda\}_{-\infty}^{\infty}$ on a $f(y_\lambda)<e^{y_\lambda}$, alors la fonction $f(x)$ est égale à Ce^x où C est une constante.

Remarquons que, dans le cas où $f(x)<e^x$, $-\infty<x<\infty$, ce résultat peut être obtenu du théorème de Liouville en se basant sur la propriété connue que la fonction $f(x)$ est régulière dans chaque domaine fini du plan des nombres complexes et que l'on a pour le module de $f(x+iy)$ l'inégalité $|f(x+iy)|\leqq f(x)$.

De cette proposition on peut tirer la suivante:

Soit $f(x)$ une fonction absolument monotone pour $x<a$ et satisfaisant pour un $b<a$ aux conditions

$$f^{(n)}(b)<K\ (n=0,1,2,3,\dots),\qquad f(x)<e^x \qquad (x\leqq b).$$

où K est une constante. Alors $f(x)$ est égale à Ce^x, où C est une constante.

En effet on a pour $x\geqq b$

$$f(x)=\sum_{n=0}^{\infty}\frac{(x-b)^n}{n!}f^{(n)}(b)<Ke^{x-b}<K_1e^x,\quad f^{(n)}(x)>0\quad (n=0,1,2,\dots)$$

et la fonction $f(x)$ est absolument monotone pour tous les x.

La proposition suivante se démontre d'une manière analogue.

Supposons que pour la fonction réelle $\varphi(x)$, deux fois dérivable pour tout x, on ait $\varphi''(x)-4\mu x\varphi'(x)+(4\mu^2x^2-2\mu)\,\varphi(x)\geqq 0$ et $\varphi(y_\lambda)<e^{\mu y_\lambda^2}$ pour une suite $\{y_\lambda\}_{-\infty}^{\infty}$ de type (9), *μ étant une constante réelle. On a alors $\varphi(x)=Ce^{\mu x^2}$, où C est une constante.*

Par la même méthode, on peut démontrer des théorèmes analogues pour les différences des fonctions. On peut généraliser aussi le théorème IV et les résultats analogues pour les fonctions satisfaisant à une inégalité différentielle.

3. Inégalités pour les suites de nombres.

Théorème V. *Soit* $\{a_n\}_1^\infty$ *une suite infinie de nombres réels et supposons qu'il existe une suite* $\{k_\lambda\}_1^\infty$ *d'entiers indéfiniment croissants et un nombre entier* $m \geqq 0$ *tels que* $a_{k_\lambda}/k_\lambda^m \to A$ *lorsque* $\lambda \to \infty$. *Désignons par*

$$\Delta a_n = a_{n+1} - a_n, \quad \Delta^2 a_n = \Delta a_{n+1} - \Delta a_n, \ldots$$

les différences de la suite $\{a_n\}$ *et supposons que pour un* $p > m$ *les différences* $\Delta^p a_n$ *soient non négatives pour* $n > n_0$. *On a alors*

$$(-1)^{p-m}(\Delta^m a_n - m!\,A) \geqq 0 \quad et \quad (-1)^{p-i}\Delta^i a_n \geqq 0 \quad (m+1 \leqq i \leqq p)$$

pour $n > n_0$. *Comme conséquence on aura pour* $n \to \infty$

$$\lim n^{-m} a_n = A, \quad \lim n^{1-m}\Delta a_n = mA, \ldots, \lim \Delta^m a_n = m!\,A,$$
$$\lim \Delta^i a_n = 0 \quad (m+1 \leqq i < p).$$

La démonstration est complètement analogue à celle du théorème I.

Théorème VI. *Soit* $\{a_n\}_{-\infty}^{+\infty}$ *une suite à deux côtés de nombres réels et supposons qu'il existe une suite d'entiers*

$$\{k_\lambda\}_{-\infty}^{+\infty}, \quad \lim_{\lambda \to -\infty} k_\lambda = -\infty, \quad \lim_{\lambda \to \infty} k_\lambda = \infty, \tag{12}$$

et un entier $m \geqq 0$ *tels que* $a_{k_\lambda}/k_\lambda^m \to 0$ *pour* $\lambda \to \pm\infty$. *Supposons encore que pour un* $p > m$ *les différences* $\Delta^p a_n$ $(-\infty < n < \infty)$ *soient non négatives. Alors* a_n *est un polynome de* n *de degré* $< m$.

Théorème VII. *Soit* $\{a_n\}_{-\infty}^{+\infty}$ *une suite de nombres réels et supposons que les différences d'ordre pair* $\Delta^{2p} a_n$ *soient non négatives pour toutes les valeurs de* n. *Supposons encore que, pour une suite* (12), $a_{k_\lambda}/k_\lambda^{2p-1} \to 0$ *pour* $\lambda \to \pm\infty$. *Alors* a_n *est un polynome de* n *de degré* $\leqq 2p-2$.

Théorème VIII. *Soit* $\{a_n\}_{-\infty}^{\infty}$ *une suite de nombres réels telle que pour un nombre pair* $2p$ *et pour tous les* n *on ait* $\Delta^{2p}(q^{-n} a_n) \geqq 0$, *où* q *est un nombre positif arbitraire. Supposons encore que pour une suite* (12) *on ait* $a_{k_\lambda} < P(k_\lambda)\, q^{k_\lambda}$ *où* $P(x)$ *est un polynome réel de degré* $2p-2$. *Alors* $a_n = Q(n) q^n$ *où* $Q(n)$ *est un polynome de degré* $\leqq 2p-2$.

Ce théorème découle immédiatement du précédent en l'appliquant à la suite $q^{-n} a_n$.

(Reçu le 3 janvier 1950)

Elementarer Beweis und Verallgemeinerung einer Reziprozitätsformel von Dedekind.

Von L. RÉDEI in Szeged.

Wir setzen

$$(1) \qquad F_m(x) = \frac{x^m - 1}{x - 1} = x^{m-1} + \ldots + x + 1 \qquad (m > 1).$$

Nachher bezeichnen $m, n\,(>1)$ teilerfremde natürliche Zahlen und m', n' eine ganzzahlige Lösung von

$$(2) \qquad mm' + nn' = 1.$$

Auch die Polynome $F_m(x)$, $F_n(x)$ sind teilerfremd, und so ist die Gleichung

$$(3) \qquad F_m(x)\, X_{mn}(x) + F_n(x)\, X_{nm}(x) = 1$$

in Polynomen $X_{mn}(x)$, $X_{nm}(x)$ lösbar. Der Grad dieser Polynome kann man kleiner als $n-1$ bzw. $m-1$ wählen, was wir im folgenden tun wollen, und dann ist die Lösung von (3) eindeutig definiert. Meines Wissens hat man sich mit der expliziten Auflösung von (3) bisher noch nicht beschäftigt, obwohl die Wichtigkeit dieser Aufgabe nicht abzusprechen ist. Ich habe vor einigen Jahren die gesagte Lösung von (3) bestimmt, sie aber noch nicht veröffentlicht. Sie lautet:

$$(4) \qquad X_{mn}(x) = \sum_{k=0}^{n-1} \left(-\left\{\frac{km'}{n}\right\} + \left\{\frac{(k-1)m'}{n}\right\} + \left\{\frac{-m'}{n}\right\} - \left\{\frac{-2m'}{n}\right\} \right) x^k,$$

wobei

$$(5) \qquad \{z\} = z - [z] - \frac{1}{2}$$

und $[z]$ die größte ganze Zahl $\leqq z$ bezeichnet. Der Grad von (4) ist wirklich kleiner als $n-1$, denn der Summand verschwindet für $k = n-1$. Selbstverständlich hat man $X_{nm}(x)$ so zu bilden, daß man in (4) nicht nur m und n miteinander sondern auch m' mit n' vertauscht.)

Es wäre unschwer die Richtigkeit der Lösungsformel (4) durch Einsetzung in (5) auszuweisen. Ich nehme hier davon Abstand und führe den Beweis in einer anderen Arbeit aus. ((4) bleibt gültig und wird etwas vereinfacht, wenn man [ ] für { } schreibt und rechts mit -1 multipliziert, aber wir behalten (4) in der angeschriebenen Form.)

Aus der durch (3), (4) angegebene Reziprozitätsbeziehung lassen sich verschiedene weitere Folgerungen ziehen. Insbesondere entsteht aus ihr der quadratische Reziprozitätssatz (für Primzahlen m, n), so daß man $x=-1$ einsetzt, was ich hier ebenfalls nicht ausführe.

Eine Reziprozitätsformel von DEDEKIND[1]) lautet so:

$$S_{mn}+S_{nm}=\frac{1}{12mn}(m^2-3mn+n^2+1), \tag{6}$$

wobei

$$S_{mn}=\sum_{k=1}^{n-1}\left\{\frac{k}{n}\right\}\left\{\frac{mk}{n}\right\}. \tag{7}$$

Außer dem Originalbeweis von Dedekind hat RADEMACHER[2])[3]) früher zwei Beweise für (6) mitgeteilt[4]), ferner teilte er in diesem Band einen weiteren recht kurzen Beweis mit[5]). Ich beweise hier (6) ebenfalls sehr leicht und elementar. Mein Verfahren besteht im wesentlichen daraus, daß ich in (3) $x=1+t$ einsetze und beiderseits die Koeffizienten von t^2 miteinander vergleiche. (Auf ähnlichem Wege ließe sich eine Fülle von Reziprozitätsformeln von der Art (6) gewinnen, weshelb wir (3) eine Verallgemeinerung von (6) ansehen können.)

Wir bemerken zunächst, daß $\left\{\frac{k}{n}\right\}$ bei festem n nur von der Restklasse $k \pmod n$ abhängt. Deshalb darf im Summand von (7) k durch $m'k$ ersetzt werden, und so geht (7) bei Berücksichtigung von (2) in

$$S_{mn}=\sum_{k=1}^{n-1}\left\{\frac{m'k}{n}\right\}\left\{\frac{k}{n}\right\} \tag{8}$$

über. Ferner bemerken wir, daß aus (5) wegen

$$\left\{\frac{k}{n}\right\}=\frac{k}{n}-\frac{1}{2} \qquad (k=1,\ldots,n-1) \tag{9}$$

sofort

$$\sum_{k=1}^{n-1}\left\{\frac{k}{n}\right\}=0 \tag{10}$$

---

1) R. DEDEKIND, *Gesammelte Math. Werke*, Bd. 1 (1930), S. 159—172.

2) H. RADEMACHER, Zur Theorie der Modulfunktionen, *Journal für die reine und angewandte Math.*, **167** (1932), S. 312—336. Hier auf Seiten 318—321 findet sin ein Gitterpunktbeweis.

3) H. RADEMACHER, Über eine Reziprozifätsformel aus der Theorie der Modulfunktionen, *Mat. és Fiz. Lapok*, **40** (1933), S. 24—34. (Ungarisch mit deutschem Auszug.)

4) Vgl. L. RÉDEI, Bemerkung zur vorstehenden Arbeit des Herrn H. Rademacher, *Ebenda*, S. 35—39. (Ungarisch mit deutschem Auszug.)

5) H. RADEMACHER, Die Reziprozitätsformel für Dedekindsche Summen, *diese Acta*, **12 B** (1950), S. 57—60.

folgt. Hier darf der Zähler k durch $m'k$ ersetzt werden, und so folgt aus (8), (9)

$$S_{mn} = \frac{1}{n} \sum_{k=1}^{n-1} \left\{ \frac{m'k}{n} \right\} k. \tag{11}$$

Endlich schicken wir noch die Formeln

$$\left\{ \frac{-m'}{n} \right\} + \left\{ \frac{-n'}{m} \right\} = -\frac{1}{mn}, \tag{12}$$

$$\left\{ \frac{-2m'}{n} \right\} + \left\{ \frac{-2n'}{m} \right\} = -\frac{2}{mn} \tag{13}$$

voran. Um diese zu zeigen, schreiben wir (2) in der Form

$$-\frac{m'}{n} - \frac{n'}{m} = -\frac{1}{mn}. \tag{14}$$

Hieraus folgt zunächst

$$\left[\frac{-m'}{n} \right] + \left[\frac{-n'}{m} \right] = -1, \quad \left[\frac{-2m'}{n} \right] + \left[\frac{-2n'}{m} \right] = -1,$$

woraus und aus (5), (14) sofort (12), (13) entsteht.

Um nunmehr (6) zu beweisen, führen wir den Operator $\mathfrak{S}$ definiert durch

$$\mathfrak{S} f(m, n, m', n') = f(m, n, m', n') + f(n, m, n', m')$$

ein. Mit dieser Symbolik schreibt sich (3) so:

$$\mathfrak{S} F_m(x) X_{mn}(x) = 1.$$

Setzt man hier (4) ein, so folgt aus (12), (13) wegen (1)

$$\mathfrak{S} F_m(x) \left(-\sum_{k=0}^{n-1} \left\{ \frac{km'}{n} \right\} x^k + \sum_{k=0}^{n-1} \left\{ \frac{(k-1)m'}{n} \right\} x^k \right) = 1 - \frac{1}{mn} F_m(x) F_n(x). \tag{15}$$

Die zweite Summe schreibt sich (wegen (5)) so:

$$\sum_{k=-1}^{n-2} \left\{ \frac{km'}{n} \right\} x^{k+1} = \sum_{k=0}^{n-1} \left\{ \frac{km'}{n} \right\} x^{k+1} + \left\{ \frac{-m'}{n} \right\} - \left\{ \frac{-m'}{n} \right\} x^n$$

und so folgt aus (15), (12), (1):

$$\mathfrak{S} F_m(x)(x-1) \sum_{k=0}^{n-1} \left\{ \frac{km'}{n} \right\} x^k = 1 - \frac{1}{mn} F_m(x) F_n(x) - \frac{1}{mn} (x-1) F_m(x) F_n(x).$$

Wieder nach (1) gilt also

$$\mathfrak{S}(x^m - 1) \sum_{k=1}^{n-1} \left\{ \frac{m'k}{n} \right\} x^k = 1 - \frac{x}{mn} F_m(x) F_n(x) + \frac{1}{2}(x^m - 1) + \frac{1}{2}(x^n - 1),$$

wobei man das zu $k=0$ gehörige Glied auf die rechte Seite geschafft hat. Setze man hier $x = 1 + t$ ein (beachte (1)) und vergleiche die Koeffizienten von t^2 miteinander:

$$\mathfrak{S}\left(m \sum_{k=1}^{n-1} \left\{\frac{m' k}{n}\right\} k + \binom{m}{2} \sum_{k=1}^{n-1} \left\{\frac{m' k}{n}\right\}\right) =$$

$$= -\frac{1}{mn}\left(m\binom{n}{3} + \binom{m}{2}\binom{n}{2} + \binom{m}{3} n + m\binom{n}{2} + \binom{m}{2} n\right) + \frac{1}{2}\binom{m}{2} + \frac{1}{2}\binom{n}{2}.$$

Die zweite Summe verschwindet, und so gilt nach (11)

$$mn\,\mathfrak{S} S_{mn} = -\left(\frac{1}{3}\binom{n-1}{2} + \frac{(m-1)(n-1)}{4} + \frac{1}{3}\binom{m-1}{2} + \frac{n-1}{2} + \frac{m-1}{2}\right) + \\ + \frac{1}{2}\binom{m}{2} + \frac{1}{2}\binom{n}{2}.$$

Die rechte Seite berechnet sich zu

$$\frac{1}{12}(m^2 - 3mn + n^2 + 1).$$

Das beendet den Beweis von (6).

(Eingegangen am 20. Januar 1950.)

Erratum.

In the paper of J. F. Koksma and R. Salem, page 88, line 4,

read $\dfrac{1}{N^2}$ *instead of* $\dfrac{1}{N}$.

INDEX. — TARTALOM.

ACTA SCIENTIARUM MATHEMATICARUM
SZEGED (HUNGARIA), ADY-TÉR 1.

TOMUM IUBILAREM
COMPLURIBUS ADIUVANTIBUS
REDIGEBAT

BÉLA SZ.-NAGY

AZ ÜNNEPI KÖTETET
TÖBBEK KÖZREMÜKÖDÉSÉVEL
SZERKESZTETTE

SZŐKEFALVI-NAGY BÉLA

Délmagyarország Nyomda NV. Szeged.

Felelős vezető: KONCZ LÁSZLÓ